Image reconstruction of remotely-sensed planetary data

Jack Wilson

Abstract

Planetary data measured from orbiting spacecraft are defined on a sphere. Therefore, the geometry of the sphere must be taken into account in their analysis. In this thesis, I develop a Bayesian image reconstruction technique, based on pixon image reconstruction, but allowing the solution to tailor its fit to structure in the image. Unlike older implementations this new method is capable of reconstructing data defined across the entire sphere without the need to partition the surface into approximately flat regions.

The thorium line data from the Lunar Prospectors Gamma Ray Spectrometer are enhanced using pixon reconstruction. The enhancement shows that at the Compoton-Belkovich Volcanic Complex, thorium is distributed over a larger (40 km $\times$ 75 km) area than the (25 km $\times$ 35 km) high albedo region normally taken to define the feature. The thorium must have been deposited during the formation of the volcanic complex, because subsequent lateral transport mechanisms, such as small impacts, are shown to be unable to move sufficient material. The morphology of the feature is consistent with pyroclastic dispersal and we conclude that the present distribution of thorium was likely created by the explosive eruption of silicic magma.

Application to Mars Odyssey Neutron Spectrometer epithermal data yields an improvement in resolution from 520 km to 290 km. This new high-resolution data is used to infer the global distribution of Water Equivalent Hydrogen (WEH) and reveals WEH-rich features near the equator at the Medusae Fossae Formation and the Tharsis Montes, which may indicate the presence of water ice. The data are also used to infer the polar abundance of CO_2 throughout the Martian year, which is found to exhibit behaviour not predicted by the Martian general circulation models.

Image reconstruction of remotely-sensed planetary data

Jack Wilson

A thesis presented in accordance with the regulations for admittance to the degree of Doctor of Philosophy

Institute for Computational Cosmology
Department of Physics
University of Durham
United Kingdom

August 2016

Dedicated to

my parents.

Declaration

The work described in this thesis was undertaken between October 2012 and July 2016 while the author was a research student under the supervision of Vincent Eke and Richard Massey in the Department of Physics at the University of Durham. Throughout this dissertation the plural pronoun 'we' is used for stylistic reasons and should be taken to refer to either the singular author, the reader and the author or, when stated explicitly, the author and collaborators. The form assumed should be apparent from the context. No part of this thesis has been submitted for any other degree at the University of Durham or any other university.

Chapter 3 of this work has appeared as Wilson et al. (2015). All of data analysis was carried out by the author and all of the text has been written by the author.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Determining the surface composition of the terrestrial planets is important for two main reasons. Firstly, knowing the bulk elemental abundances reveals the planets' history. Relative abundances between planets provide evidence of the details of their formation from the protoplanetary disk and subsequent interactions (Righter & O'Brien, 2011), and the concentration of radioactive elements such as thorium allows their thermal history to be inferred. Secondly, future utilization of planetary resources depends on knowing their locations and abundances. The location of the most concentrated and accesible deposits of near-surface hydrogen and water remain an important questions for future manned exploration of the Moon and Mars (Squyres, 2011).

There exist many techniques for remotely determining surface the composition of planets (Pieters & Englert, 1993). These include: visible and infrared imaging and spectroscopy which are used to infer the mineralogical composition of the top few μm of surfaces; radar sounding, which reveals density contrasts on the scales of tens of cm to m to depths of several km (Picardi et al., 2004; Seu et al., 2007); and neutron and gamma ray spectroscopy, which are used to measure bulk elemental abundances in the top few tens of cm (Feldman et al., 1999; Boynton et al., 2004).

Although orbital gamma ray and neutron spectroscopy have become a standard remote sensing tool for measuring planetary composition, the large spatial footprint of the detectors represents a significant drawback of this technique. This broad spatial resolution is due to that fact that nearly all orbital gamma ray and neutron measurements use omni-directional detectors for which the full width half maximum of the spatial response function is $\sim$ 1-1.5 times the spacecraft's height above the planetary surface. For the Lunar Prospector and Mars Odyssey missions, the optimum spatial resolutions were 54 and 600 km (Maurice et al., 2004; Prettyman et al., 2004), respectively. Consequently, many

important science questions cannot be addressed without improving on current data. One way to do this is with new missions orbiting at low altitudes that provide better resolution data. However, this is necessarily expensive and, in the case of planets with substantial atmospheres, may not be technically practical. A more economical and immediately available alternative to new missions is to use the statistical techniques of image reconstruction to improve the data that are already available. It is this approach that we will explore in this thesis, in which image reconstruction techniques will be developed for use on the sphere and applied to remotely sensed gamma ray and neutron data to better understand the surface composition of the Moon and Mars.

1.1 A brief history of determining planetary surface composition

In the following sections we will consider separately the development of knowledge of the surface composition of the Moon and Mars before detailing the techniques of gamma ray and neutron spectroscopy, from which the data used in this thesis are obtained. Our progress in understanding will be seen to follow a similar path for both bodies, with both known since antiquity, understood further following the invention of the telescope and the use of earth based imaging techniques and finally greatly expanded with the advent of landers and orbiting missions starting in the mid-20th century. This section is necessarily not exhaustive and will focus on remote sensing of the surface neglecting landers and rovers as well as instruments and missions designed to probe the atmospheres or exospheres of planets.

1.1.1 The Moon

Speculation of a quasi-scientific nature about the composition of the lunar surface dates back to at least the ancient Greeks, with both Plutarch and Democritus speculating on the possibility of the lunar dark spots being similar to the Earth's oceans (Whitaker, 1999). This theory was believed, in one form or another, into the early modern period when, around 1600, William Gilbert made the first systematic study of the features on the

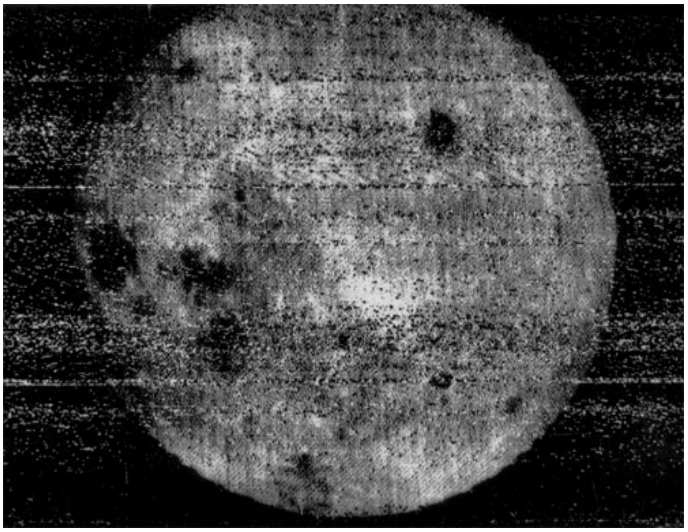

Figure 1.1: The first image of the farside of the Moon returned by Luna 3 (Shelton, 1968).

Moon's surface (Gilbert, 1651). However, Gilbert believed (as did Da Vinci and Kepler) that the dark regions were islands surrounded by a brighter ocean.

Early telescopic observations revealed the bright lunar highlands to be heavily cratered terrain (Galilei, 1610). Galilei (1610) was circumspect in his description, avoiding explicitly labelling the dark areas as seas. However, this seems to be how many interpreted his work and when atlases of lunar geography came to be created in the eighteenth century, the dark patches are uniformly labelled 'mare', 'lacus' or 'oceanus' (Schröter, 1791).

Later telescopic observations in the late nineteenth and early twentieth centuries allowed for the characterization of the geologic character of the Moon's nearside at a spatial resolution up to 1 km (Gilbert, 1893; Spurr, 1944; Baldwin, 1949). The stark differences between the bright, mountainous highlands and the smooth, dark maria were finally attributed to flooding of depressions by low-viscosity lavas. The presence of large numbers of circular craters had been known since the earliest telescopic observations. Initial they were thought to be volcanic in origin, but in the twentieth century opinion increasingly began to favour an impact origin. Work by Shoemaker (1963) demonstrated the role of cratering on Earth, and the first spacecraft-based lunar observations supported the interpretation that the Moon's surface processes are dominated by impact events.

The first, entirely successful, lunar mission was the USSR's Luna 2, which discovered time variation in the energy spectrum of the Earth's Van Allen belt and showed that the

Moon had no similar belt. Luna 3 followed less than one month later on 6th October 1959 (Shelton, 1968) and returned the first pictures of the far side of the Moon. This image, shown in figure 1.1, revealed surprising compositional asymmetry between the lunar near and far sides, with the maria covering 30% of the nearside but only 1% of the farside (Shearer et al., 2006).

A flotilla of American and Russian flyby, impactor, orbiter and lander missions followed (Spudis, 1999), including Luna 10, the first lunar orbiter, that carried a gamma ray spectrometer (Vinogradov et al., 1966) and the five spacecraft in NASA's Lunar Orbiter program that mapped 99% of the Moon's surface at a resolution better than 60 m (Byers, 1977). These missions culminated in six Apollo missions that involved manned landing and, importantly for remote sensing, the return of large volumes of samples from known locations on the surface. These samples have been used to calibrate remotely sensed data by providing known ground truths (Shkuratov et al., 2001). After the final Apollo mission in 1972, several Luna landers, performing sample return, landed before a 15 year hiatus in lunar missions occurred.

The pause in lunar missions was ended in 1994 when Clementine, co-developed by NASA and the Ballistic Missile Defense Organization, became the first of a new generation of multi-instrumented spacecraft to be placed in orbit around the moon (Nozette et al., 1994) (though the Japanese Hiten had been intentionally crashed into the lunar surface in 1993 it was primarily an engineering test and returned no data about lunar surface composition (Burnham, 1993; Uesugi, 1996)). Clementine contained four cameras used for remote sensing each with a resolution of decametres to kilometres: the Ultraviolet/Visible camera, which was designed to operate at five different wavelengths and used to infer global FeO and TiO_2 surface abundances (Blewett et al., 1997); a near-infrared CCD camera, which suffered from calibration problems, but was used to successfully map olivine in the Aristarchus and Copernicus craters (McEwen et al., 1994; Le Mouélic et al., 1999; Daydou et al., 2003); a Long Wavelength Infrared Camera; and a high-resolution camera (HIRES), which returned (at the time) the highest resolution spectral images of the Moon (Robinson et al., 2003). The Clementine mission showed that the apollo samples did not represent the lunar crust as a whole, and revealed complex crustal stratigraphy (Robinson & Riner, 2005).

Clementine was followed, four years later, by NASA's Lunar Prospector (Binder, 1998). This orbiter carried gamma ray (GRS), neutron and alpha particle spectrometers (Feldman et al., 2004). The GRS produced the first global gamma ray flux data set from the lunar surface, which were used to infer global elemental maps of Al, Fe, K, O, Si, Ti and Th (Lawrence et al., 1998a). The neutron spectrometer was able to identify polar hydrogen deposits (Feldman et al., 2000), which in combination with topographic data, identifying permanently shaded craters, and image reconstruction, implied that the excess H is concentrated within the permanently shaded craters (Eke et al., 2009).

In 2007 both China's first mission, Chang'E 1 (Ouyang et al., 2008), and Japan's first science mission, the SELenological and ENgineering Explorer, SELENE, successfully entered lunar orbit (Kato et al., 2008). SELENE, more often called Kaguya, contained 15 instruments including X-ray and and gamma ray spectrometers to measure elemental composition; multiband and spectral imagers, measuring mineralogical distribution; and a radar sounder and laser altimeter to take topographic measurements (Kato et al., 2008). Chang'E 1 is the first lunar probe to conduct passive, multi-channel, microwave remote sensing of the Moon by using a microwave radiator. These were joined in 2008 by India's first lunar orbiter Chandrayaan-1. Chandrayaan-1 contained multiple UV/Visble/IR imagers and spectrometers (Bhandari, 2005) including the Moon Mineralogy Mapper, M^3, which successfully detected the spectral signature of hydrated minerals on the Moon (Bhattacharya et al., 2013). The spacecraft also contained an active synthetic aperture radar system, Mini-SAR, designed to search for lunar polar water ice within permanently shadowed craters.

NASA's most recent orbiter whose focus is surface mapping is the Lunar Reconnaissance Orbiter (LRO), which was launched in 2009 and remains operational today (Vondrak et al., 2010). The LRO contains seven instruments the: Cosmic Ray Telescope for the Effects of Radiation, Diviner Lunar Radiometer Experiment, Lyman-Alpha Mapping Project, Lunar Exploration Neutron Detector (LEND), Lunar Orbiter Laser Altimeter, Lunar Reconnaissance Orbiter Camera (LROC) and synthetic aperture radar instrument (Mini-RF).

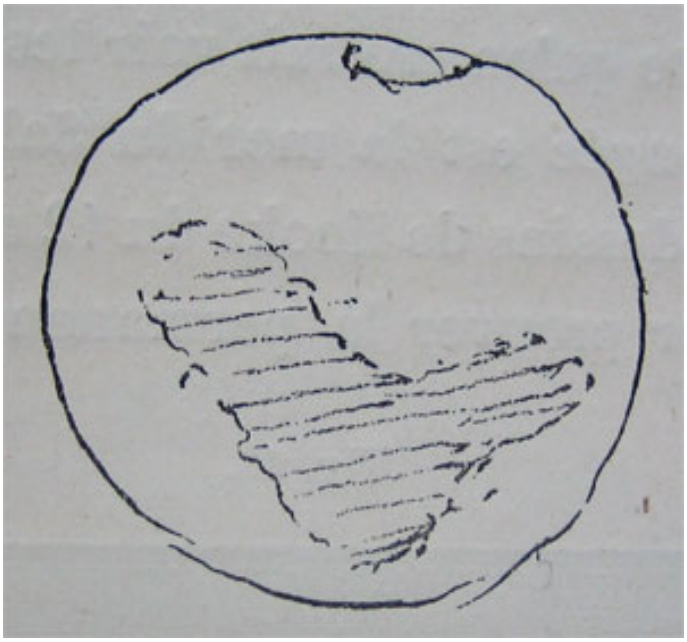

Figure 1.2: Huygens' sketch of his observation of Mars made on 13th August 1672, south is at the top of the image. The south polar cap is seen at the top of the image along with an albedo feature that may be Syrtis Major.

1.1.2 Mars

Unlike the Moon, the first hints at compositional variation on Mars required observation through a telescope. Cassini (1666) noted the presence of of the bright polar cap at the Martian north pole, before Huygens in 1672 identified its counterpart in the south (Huygens, 1698). Huygens sketch made during this observation is reproduced in figure 1.2. Both Cassini and Huygens noted the presence of other albedo features and that their shape varied little on the time scale of hours, thus they were able to use their observations to estimate the Martian day to within 1 % of its currently accepted value. Longer term changes were suggested to be linked to an annual water cycle. However, it wasn't until Herschel (1784) that it was suggested that the polar caps were water ice sheets, like those observed on Earth, that advance and retreat annually.

In the nineteenth and early twentieth centuries a recurring phenomenon was noted by several observers (Fournier, 1914), where darkening of the surface spreads north from the south pole during spring. Motivated by the green tint that was seen in the darker areas (Kuiper, 1955), early explanations for this observation favoured spreading vegetation, before further observation suggested that this colouration was an observational effect

(Kuiper, 1957). Finally dust storm activity was suggested by Sagan & Pollack (1969) as an explanation for periodic darkening of the near-polar regions, which is now known to be correct.

Understanding greatly changed with the first successful flyby of the planet Mars by the Mariner 4 spacecraft on 14th and 15th July 1965. The 22 images captured by its equipped television camera were the first detailed pictures of the Martian surface and revealed a heavily cratered world lacking obvious signs of past or present water activity (Anderson, 1965). Mariners 6 and 7 arrived four years later and, surveying a larger part of the surface, were able to find features appearing fluvial in origin including smooth plains in Hellas Basin and gullies. The infrared radiometers carried by both of the spaceraft showed the polar caps to be made of CO_2, and not water ice. Imagers on subsequent orbiters and landers in the 1960s and early 1970s pointed to the existence of a warmer, wetter period early in Mars' history. This was followed by a 20 year pause in the robotic exploration of Mars (Taylor, 2010).

In late 1997, the Mars Global Surveyor (MGS; Albee et al., 2001), the first in a series of modern Mars missions carrying new kinds of instruments, arrived at Mars. In addition to a higher resolution camera, the Mars Orbiter Camera (Malin et al., 2010), it carried the Thermal Emission Spectrometer (TES Christensen et al., 2001) a hyperspectral thermal infrared (IR) spectrometer sensitive from 6 to 50 μm, and visible-NIR (0.3 to 2.9 μm) bolometer. TES was used to determine global surface composition by comparing returned IR spectra with those measured on Earth (Bandfield et al., 2000). The ancient southern highlands were found to be basaltic whereas the northern lowlands were predominantly andesitic (or covered by weathered basalt dust). Grey, crystalline hematite was identified in some locations, suggesting the existence of prolonged wet periods on the Martian surface (Christensen et al., 2001).

MGS was followed by NASA's 2001 Mars Odyssey, which entered orbit in October 2001 and remains operational today. Thermal Emission Imaging System (THEMIS) a visible/IR camera. Like TES, THEMIS can be used to infer the mineralogy of the Martian surface, but it does so on much finer spatial scales with a resolution of 100 m compared to TES' 3 km (Christensen et al., 2004). Odyssey also contained the Gamma Ray Suite of instruments, consisting of a gamma ray spectrometer, neutron spectrometer, and high

energy neutron detector (the neutron detector is described in more detail in the following section as its data will be used in chapters 4, 5 and 6 of this thesis; Boynton et al., 2004). Unlike visible/IR spectrometers which see only the top few μm, the Gamma Ray Suite's instruments are sensitive to the composition of the top few 10s of cm of the surface. Additionally, they are sensitive to the bulk atomic composition rather than its chemical state. These instruments have been used to create maps of elements, including Si, Cl, K, Fe, and Th, in the low- and mid-latitude regions of Mars (Boynton et al., 2004), in addition to revealing the presence of substantial subsurface hydrogen, likely in the form of water ice, within 30° of the poles (Feldman et al., 2002).

The European Space Agency's Mars Express was launched in June and arrived in orbit December 2003. It contains a camera, the High Resolution Stereo Camera, capable of 2 m/pixel resolution (Gwinner et al., 2016); a visible and infrared mapping spectrometer, the Observatoire pour la Minéralogie, l'Eau, les Glaces et l'Activité (OMEGA) (Bibring et al., 2006); and, for the first time on a Mars-orbiting craft, a subsurface sounding radar altimeter, Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) (Nielsen, 2004). MARSIS is sensitive to depths of up to four kilometres in the ice-rich polar layered deposits and to depths of up to several hundred meters over the rest of the Martian surface (Picardi et al., 2005). The instrument has been used to create maps of the dielectric constant across Mars, these maps imply the existence of ice rich material coincident with the near-polar region identified in gamma ray and neutron data as being abundant in hydrogen nuclei (Mouginot et al., 2012).

In 2006 the Mars Reconnaissance Orbiter entered orbit around Mars (Vondrak et al., 2010). It carried a range of cameras with different resolutions and fields of view (Vondrak et al., 2010); a visible/IR spectrometer, CRISM; and a subsurface sounding radar, the SHAllow RADar (SHARAD) sounder. SHARAD operates at a higher frequency than MARSIS and penetrates to less depth but achieves a finer resolution (Seu et al., 2007).

1.2 Gamma ray and neutron spectroscopy

Neutron and gamma ray spectroscopy have been used extensively to determine the elemental composition of the surfaces of planetary bodies with little or no atmosphere,

notably the Moon and Mars (Elphic et al., 1998; Lawrence et al., 1998b; Boynton et al., 2004). The detected neutrons and gamma rays are the result of cosmic rays interacting with planetary surfaces and atmospheres or by the naturally occurring decay of radioisotopes within the surface (Reedy, 1978). The resultant flux of gamma rays and neutrons encodes information about the distribution and abundance of various major and minor elements and can be measured from orbit, thus enabling global mapping of elemental composition. Given the mean free path length of neutrons, the radiation that escapes into space typically originates from less than 1 m of the surface. This fact provides gamma ray and neutron spectroscopy a major advantage over other remote sensing methods in their ability to accurately determine elemental abundance at depth, not merely of the immediate surface. The price for this is the relatively poor spatial resolution achieved by neutron and gamma ray spectrometers, which achieve resolution comparable to the spacecraft orbital distance compared to the few kilometre to metre resolution achievable with imaging and spectroscopy (Pieters & Englert, 1993). This is due to the fact that the spectrometers are, typically, insensitive to incidence direction. Therefore higher resolution can be achieved by reducing the height at wich the spacecraft orbits. Low altitude is also desired as the production rate of neutrons and gamma rays is low (Lingenfelter et al., 1961; Reedy et al., 1973) compared with photons in the visible range. So, unlike optical techniques, distances closer than a few hundred kilometers are needed in order to obtain a strong signal. Additionally, measurements of the surface of planets with dense atmospheres, such as the Earth and Venus, is not achievable, though variations in atmospheric composition can be measured. The rest of this section will explain the principles behind neutron and gamma ray spectroscopy in more detail, focussing on the spectrometers from which the data used later in this thesis was obtained.

1.2.1 Production of gamma rays and neutrons

Neutrons and gamma rays (along with a host of other short-lived particles) are produced by the interaction of energetic particles and cosmic rays with planetary surfaces and atmospheres. For the purposes of gamma ray and neutron spectroscopy it is the interaction of high energy Galactic Cosmic Rays (GCRs) with planetary surfaces and atmospheres that are of interest, as they give rise to the bulk of the particles that can be detected from

orbit (Lingenfelter et al., 1961). Gamma rays are also continually produced in the decay of radioactive elements such as K, Th, and U (Lucey et al., 2006).

Around 90% of GCRs consist of individual protons, with an average flux in the inner solar system of $\sim$4 $cm^{-2}s^{-1}$ and a broad energy distribution reaching many GeV (Gaisser, 1990). Although their flux and energy distribution is approximately constant with time, the solar cycle is known to affect the number and energy of GCRs reaching the inner solar system. At $\sim$94 AU from the Sun, the solar wind undergoes transitions from supersonic to subsonic speeds, in the termination shock. The region between the this shock and the heliopause (the boundary between the region of space dominated by the Sun and the interstellar medium) acts as a barrier to GCRs, decreasing their flux at energies $\leq$1 GeV by $\sim$90% (Parker, 1965; Langner et al., 2003). Thus, higher fluxes of GCRs are observed during periods of low solar wind activity. Additionally, more low-energy GCRs penetrate the heliosphere during periods of low solar activity, resulting in a shift in the GCR energy distribution towards lower energy particles. The flux and energy distribution of GCRs are important factors in determining the energy distribution, rate and depth of origin of neutrons and gamma rays detected in orbit (Lingenfelter et al., 1961).

The GeV-scale kinetic energy of GCRs is large compared with the binding energy of nuclei, for example the binding energy per nucleon of an ^{56}Fe nucleus, which is one of the most tightly bound nuclei, is 8.8 MeV (Wapstra & Audi, 1985). Therefore, the collision of a high energy GCR with a nucleus in a planetary surface or atmosphere results in the break up of the nucleus with the energy of the incident particle transferred to the nucleons. This results in the emission of secondary particles, including protons and neutrons, via spallation followed by de-excitation of the remaining nuclei. These secondary particles will undergo further collisions with nuclei until the initial energy of the cosmic ray is absorbed by the medium. Since most of the gamma ray production is caused by reactions with neutrons, we will focus our attention on how neutrons slow down in matter (Reedy et al., 1973).

Neutrons, produced in spallation reactions, undergo multiple interactions with nuclei, transferring their energy to the surface, or atmosphere, before eventually being absorbed or escaping to space, this process is known as 'moderation'. There are three types of collision that are of interest in gamma ray and neutron spectroscopy: inelastic collision,

in which the incident neutron is absorbed, forms a compound nucleus, which decays by emitting one or more neutrons followed by the emission of gamma rays; elastic scattering, in which the colliding particles bounce off one another, conserving kinetic energy; and thermal neutron capture, in which a low energy neutron is absorbed and one or more gamma rays are emitted. The gamma rays produced in these de-excitation processes have discrete energies, so the energy and intensity of gamma ray lines can be used to characterize the elemental composition of the medium.

The neutrons produced in GCR spallation processes have energies from $\sim 10^{-4}$ to 10^{10} MeV (Feldman et al., 1998). Within this range the energy distribution depends on the elemental composition of the surface and atmosphere, which determines the cross-sections for the various possible neutron interactions. In practice the energy range is split into three regimes (illustrated in figure 1.3) that each contain compositional information. These regimes are: thermal neutrons, which have undergone a large number of collisions, have energies less than ~ 0.4 eV and are in thermal equilibrium with the surface; epithermal neutrons, which have energies above thermal neutrons but less than 0.6 MeV; and fast neutrons, which have undergone few, or no, elastic collisions and have energy above the epithermal neutrons. Fast neutrons have undergone few collisions so their flux is dependent on the mean atomic mass of the soil as larger nuclei produce more neutrons in spallation reactions (Gasnault et al., 2001). On the Moon and Mars the variation in fast neutron flux is dominated by variation in Fe abundance (Maurice et al., 2000). The importance of elastic scattering is seen by considering the maximum energy loss during such a collision,

$$L = E\left(1 - \left(\frac{A-1}{A+1}\right)^2\right), \tag{1.2.1}$$

where E is the pre-collision neutron energy in the target's frame and A is the atomic mass of the target nucleus (Prettyman, 2007). Therefore, on collision with a H nucleus, which has roughly the same mass as a neutron (i.e. $A \approx 1$), a neutron could lose all of its energy. Whereas for a collision with Fe, the maximum energy loss is 7%. Consequently, for media with a high H abundance, such as water ice, energy loss by elastic collisions is high and neutrons slow down more quickly than for materials that do not contain H. Thus the flux of epithermal neutrons reflects the abundance of H (Lingenfelter et al., 1961; Feldman et al.,

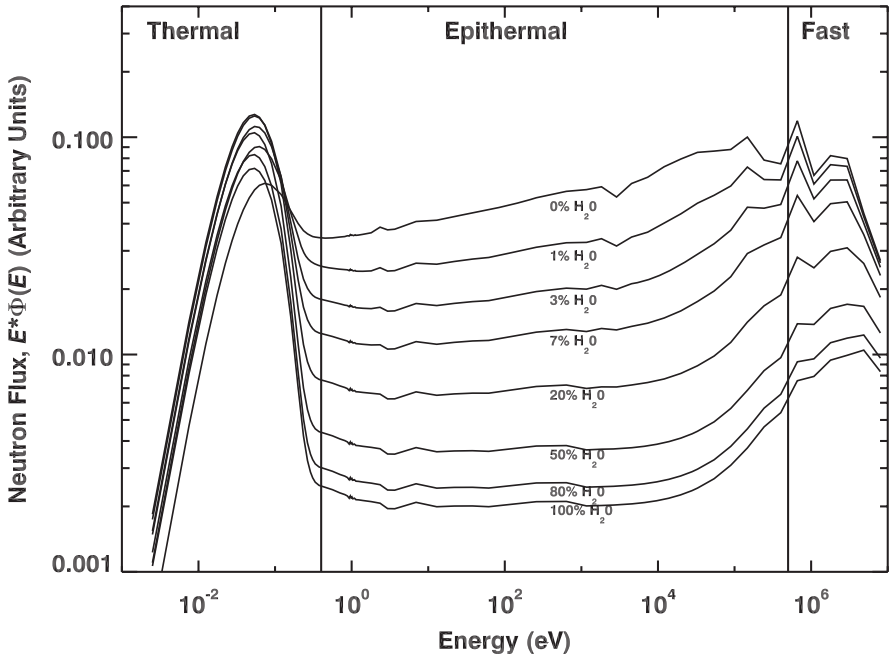

Figure 1.3: Modelled neutron flux, $\Phi(E)$, for a lunar soil with various amounts of added water. The fluxes are plotted as flux times energy, $E\Phi(E)$, to better show the variation of epithermal neutrons with increasing water content. This figure is adapted from Lawrence et al. (2010).

1998, 2001). Absorption of neutrons by thermal neutron capture significantly influences the population of thermal neutrons. Elements such as H, Cl, Gd, Fe, Sm and Ti have relatively high cross sections for thermal neutron capture, so the thermal neutron flux is most sensitive to their abundance (Feldman et al., 1998; Elphic et al., 1998, 2000).

1.2.2 Detection of gamma rays and neutrons

Radiation spectrometers rely on the detection of ionized particles produced within some sensitive medium due to its interaction with energetic gamma rays and neutrons. Gamma ray interactions include photoelectric absorption, Compton scattering and pair production, all of which cause the gamma ray to lose energy to the sensitive medium (Fichtel & Trombka, 1997). Neutrons, being neutral particles, do not directly ionize. However their moderation, as described in the previous section, deposits energy in the medium and creates gamma rays that may also be detected. There are several possible designs of spectrometer, including gas proportional counters (Lockwood, 1973; Feldman et al., 1999)

and semiconductor radiation detectors (Prettyman et al., 2011). However in this section we will describe the scintillator type as this is the form of detector from which the data used in this thesis are drawn.

Scintillator based radiation detectors consist of a scintillator, a material that undergoes luminescence when excited by ionizing radiation, coupled with a photon detector, either in the form of photomultiplier tubes or photodiode. The number and energy of photons produced is determined by the energy deposited into the scintillator. Therefore detecting these photons allows the energy of the ionizing particle to be found.

The three different neutron energy ranges are separated using two techniques. Both thermal and epithermal neutrons are captured by neutron absorbing nuclides in the detector, causing excitation. Subsequent de-excitation results in the release of a gamma ray with a characteristic energy. Thus, the signature of both thermal and epithermal neutrons is a single pulse that has an amplitude that is characteristic of the de-excitation of the neutron-absorbing nuclide in the detector. Low energy, thermal neutrons can be screened out by placing a cadmium foil shield on the scintillator. Cadmium strongly absorbs neutrons with energies less than 0.4 eV (Boynton et al., 2004; Feldman et al., 2004). Comparison between shielded and unshielded scintillators can be used to determine the separate contribution of thermal and epithermal neutrons. High energy neutrons undergo elastic scattering with protons in the scintillator, producing light, before they too are absorbed into the material. Fast neutrons are, therefore, characterized by a double-pulse time signature, corresponding to an initial elastic collision followed by de-excitation after absorption (Leo, 1987).

Scintillators are also used for gamma ray spectroscopy. Gamma ray detections are characterized by single peaks, the energies of which are determined by the wavelength of the incident gamma ray, which is indicative of the emitting nucleon (Feldman et al., 2004).

1.2.3 Lunar Prospector Gamma Ray Spectrometer

The Lunar Prospector spacecraft entered orbit in January 1998. It was operational in a near-polar, near-circular, 100 km altitude orbit between 17 January and 25 October 1998, before in late October 1998 the altitude was reduced to an average of 30 km (Feldman

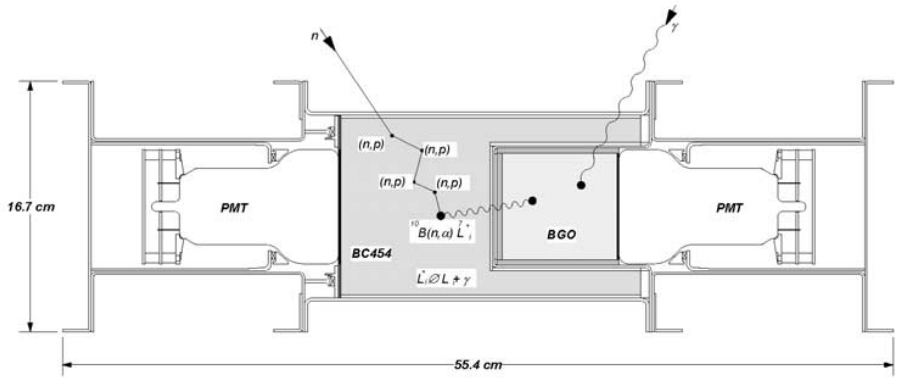

Figure 1.4: The Lunar Prospect Gamma Ray Spectrometer, showing the BGO scintillator borated-plastic anticoincidence shield (BC454) and photomultiplier tubes (PMT). This figure is adapted from Feldman et al. (2004)

et al., 2004). The mission was ended by deliberately crashing the spacecraft into the Moon near the south pole on 31 July 1999 (Goldstein et al., 1999). One of the main goals of the mission was to determine global maps of elemental composition of the lunar surface, with special emphasis on a search for polar water-ice deposits implied by the H abundance. A Gamma-Ray Spectrometer (GRS) and a Neutron Spectrometer were included onboard the spacecraft for this purpose (Binder, 1998). The Lunar Prospector mission is one of only two lunar missions to have carried a gamma ray spectrometer to the Moon, with the other being the Apollo Gamma Ray Spectrometer.

The GRS was a scintillator type detector consisting of a 7.1 cm diameter by 7.6 cm long bismuth germanate (BGO) crystal, within a 12 cm diameter by 20 cm long, borated-plastic anticoincidence shield (Feldman et al., 1999). The instrument is illustrated in figure 1.4. The anticoincidence shield was present to detect energetic charged particle events so that they could be prevented from contaminating the gamma ray signal. However, it was also capable of functioning as an additional neutron detector for the spacecraft (Feldman et al., 2004). The gamma ray energy detection range of the GRS was between 0.3 and 9 MeV with a resolution of 17.6 keV. Its accumulation time was 32 s, which corresponds to a ground track distance of about 50 km (Feldman et al., 2004).

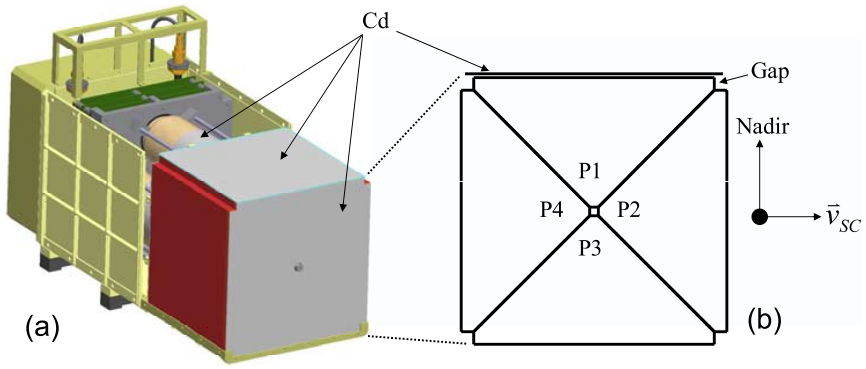

Figure 1.5: (**a**) An engineering drawing of the neutron spectrometer, which shows the placement of Cd filters around the spectrometer. (**b**) An illustration of the prisms' orientation relative to the spacecraft velocity vector ($\vec{v}_{SC}$). This figure is taken from Prettyman et al. (2009)

1.2.4 Mars Odyssey Neutron Spectrometer

The Mars Odyssey spacecraft entered orbit on 24th October 2001 and is now in a polar mapping orbit around Mars at a height of approximately 400 km. The nuclear spectroscopy payload consists of a gamma ray spectrometer, a neutron spectrometer (MONS), and high energy neutron detector (Boynton et al., 2004).

The MONS instrument consists of a boron-loaded plastic scintillator cube ($\sim$ 10 cm on a side), which has been diagonally segmented into four prisms. The prisms are optically isolated from one another and read out by separate photomultiplier tubes, illustrated in figure 1.5. Of the four prisms one is nadir facing (P1), one faces in the direction of spacecraft motion (P2), one faces zenith (P3), and the final one faces opposite the direction of spacecraft motion (P4). P1 is shielded agains thermal neutrons with Cd foil. Therefore, P1 is sensitive only to epithermal and fast neutrons originating from the surface and atmosphere (Boynton et al., 2004).

In addition to shielding, a Doppler technique can be used to separate the contribution of thermal and epithermal neutrons to the detected count rate (Feldman & Drake, 1986). To do this, MONS makes use of the orbital speed of the spacecraft relative to that of thermal neutrons. Mars Odyssey is travelling at $\sim$ 3400 ms^{-1}, which is comparable to that of a 0.05 eV neutron. Neutrons below this speed will not be seen in prism P4, thus, P4 is not sensitive to thermal neutrons. Conversely, P2 ploughs in to thermal neutrons

ahead of the spacecraft. However, both prisms have approximately the same sensitivity to epithermal neutrons. Therefore, the thermal neutron flux is given by the difference between the count rates for prisms P2 and P4 (Boynton et al., 2004).

1.3 Thesis overview

The remainder of this thesis is structured as follows:

- **Chapter 2:** we review the mathematical foundations of image reconstruction, upon which the work performed in the remainder of this thesis is based, and develop new image reconstruction techniques that will be useful for remotely-sensed planetary data.

- **Chapter 3:** we perform a pixon reconstruction of the Lunar Prospector Gamma Ray Spectrometer thorium data around the Compton-Belkovich Volcanic Complex. The geological implications of the revealed thorium distribution are discussed.

- **Chapter 4:** we characterize the properties of the MONS and apply our new image reconstruction process to the data. The achieved improvement in resolution is quantified and comparisons to surface measurements are made.

- **Chapter 5:** we discuss several regions that our reconstructions of the MONS data reveal to contain interesting structure.

- **Chapter 6:** we apply our image reconstruction technique to MONS data split temporally in order to measure seasonal changes in the Martian polar CO_2 caps. Non-polar variation in water equivalent hydrogen abundance will also be searched for.

- **Chapter 7:** we present our conclusions, summarising the work in this thesis and outlining promising avenues of future research.

CHAPTER 2

Image reconstruction and the pixon method

The later chapters of this thesis use the results of a pixon image reconstruction process applied to various remotely-sensed planetary data sets. This chapter begins by describing some mathematics that will be useful in this thesis before details of the principles behind image reconstruction in general, along with the pixon method in particular are described. It then describes improvements made to previous implementations of the pixon method, and their adaptation to the sphere, are then described.

2.1 Mathematical preliminaries

Here we will describe some mathematical functions, expansions, transforms and other operations that will be used later when implementing image reconstruction algorithms on the sphere. The details are included here to make the conventions and notations adopted explicit but the reader familiar with spherical harmonic analysis/synthesis and convolution on the sphere may wish to skip this section.

2.1.1 Legendre and associated Legendre functions

The Legendre functions, $P_l(x)$, are the solution to Legendre's differential equation:

$$\frac{d}{dx}\left((1-x^2)\frac{d}{dx}P_l(x)\right)+l(l+1)P_l(x)=0. \tag{2.1.1}$$

The solutions for integer l form a sequence of orthogonal polynomials, on the interval $x \in [-1,1]$, known as the Legendre polynomials of degree l (Riley et al., 2006). Where

orthogonality is defined for two functions f and g as $\int f(x)f(g)dx = 0$. The polynomials may be calculated using Rodrigues' formula,

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l. \tag{2.1.2}$$

The associated Legendre polynomials form a superset of these Legendre polynomials and can be defined by

$$P_l^m(x) = (-1)^m(1-x^2)^{m/2}\frac{d^m}{dx^m}P_l(x), \tag{2.1.3}$$

for $0 \leq l$, $0 \leq m \leq l$ and where the superscript m denotes the order of the function, not a power of P_l (Riley et al., 2006). Although these functions are conventionally called the Legendre polynomials, they are, in fact, polynomial only for even m, being the square root of a polynomial for odd m. In the case $m = 0$ one recovers the Legendre polynomials, i.e. $P_l^0(x) = P_l(x)$. Using Rodrigues' formula, equation (2.1.2), along with equation (2.1.3) the associated Legendre polynomials can be expressed as

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l, \tag{2.1.4}$$

which enables the extension to $-l \leq m \leq l$. Making the substitution $m \to -m$ it can be seen that the solutions with negative m are related to those with positive m by

$$P_l^{-m}(x) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(x). \tag{2.1.5}$$

2.1.2 Spherical harmonics

The spherical harmonics are the angular part of the solutions to Laplace's equation, $\nabla^2 f = 0$, i.e. the divergence of the gradient of a scalar field is zero, in spherical polar coordinates (Riley et al., 2006). Solving this equation, via separation of variables, yields the spherical harmonics as the angular part of the solution:

$$Y_l^m(\theta,\phi) = (-1)^m\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\phi}, \tag{2.1.6}$$

where θ and ϕ are colatitude and longitude, respectively, and the normalization has been chosen such that the functions are orthonormal, i.e.

$$\int\int Y_l^m(\theta,\phi)(Y_{l'}^{m'})^*(\theta,\phi)d\Omega = \delta_{ll'}\delta_{mm'} \tag{2.1.7}$$

where $d\Omega = \sin\theta d\theta d\phi$ and $\delta_{ll'}$ is the Kronecker delta.

We note here the addition theorem for spherical harmonics:

$$P_l(\boldsymbol{\xi}\cdot\boldsymbol{\xi'}) = \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y_l^m(\boldsymbol{\xi'})\,(Y_l^m)^*\,(\boldsymbol{\xi}), \tag{2.1.8}$$

where $\boldsymbol{\xi}$ labels a point on the unit sphere. This formula will be used later when calculating convolutions on the sphere.

Spherical harmonic transforms

The spherical harmonics form a complete basis of square-integrable functions on the 2-sphere, $\mathbb{S}^2$ (i.e. those functions f for which $\int_{\mathbb{S}^2}|f(\theta,\phi)|^2 d\Omega < \infty$). Thus any square-integrable function, f, can be written (Riley et al., 2006)

$$f(\boldsymbol{\xi}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} f_l^m Y_l^m(\boldsymbol{\xi}). \tag{2.1.9}$$

Orthogonality of the basis functions means that the spherical harmonic coefficients are given by

$$f_l^m = \int_{\mathbb{S}^2} f(\boldsymbol{\xi})\,(Y_l^m)^*\,(\boldsymbol{\xi})d\boldsymbol{\xi}, \tag{2.1.10}$$

where * is a complex conjugate. In this thesis only real functions will be considered. These have the symmetry relation $f_l^{-m} = (-1)^m\left(f_l^m\right)^*$, which will reduce the computation and memory required for spherical harmonic transforms.

In practice, these transforms are carried out on data defined in a grid of pixels, include contributions only up to some maximum degree $l_{\max}$ and the integration is done numerically. Typically $l_{\max} \approx \sqrt{N_{\rm pix}}$, with $N_{\rm pix}$ the number of pixels in the image. The

transforms are then written

$$f(\mathbf{x}) = \sum_{l=0}^{l=l_{\max}} \sum_{m=-l}^{l} f_l^m Y_l^m(\mathbf{x}), \tag{2.1.11}$$

$$f_l^m = \sum_{\mathbf{x}=\mathbf{0}}^{N_{\rm pix}} f(\mathbf{x}) w(\mathbf{x}) \left(Y_l^m\right)^*(\mathbf{x}), \tag{2.1.12}$$

where $\mathbf{x}$ labels a pixel and $w(\mathbf{x})$ is the weighting appropriate to pixel $\mathbf{x}$ given the pixelization. The time complexity for both of these operations is $O(l_{\max}^2 N_{\rm pix})$. However for pixel grids arranged on N_θ iso-latitude rings with pixels within each ring having the same weight and equal spacing the transforms can be written

$$f(\mathbf{x}) = \sum_{m=-l_{\max}}^{l_{\max}} \sum_{l=|m|}^{l_{\max}} f_l^m Y_l^m(\theta(\mathbf{x}),0) \exp\left(im\phi_{0,y} + \frac{2\pi imx}{N_{\phi,y}}\right), \tag{2.1.13}$$

$$f_l^m = \sum_{x=0}^{N_{\phi,y}-1} \sum_{y=0}^{N_\theta-1} f(\mathbf{x}) w(\mathbf{x}) \left(Y_l^m\right)^*(\theta(\mathbf{x}),0) \exp\left(-im\phi_{0,y} - \frac{2\pi imx}{N_{\phi,y}}\right), \tag{2.1.14}$$

where y indexes the iso-latitude rings and x the pixels within these rings, $\theta(\mathbf{x})$ is the colatitude in pixel $(\mathbf{x})$, $N_{\phi,y}$ are the number of pixels within the yth iso-latitude ring and $\phi_{0,y}$ is the longitude of the first element in ring y. Defining

$$F_{m,y} = \sum_{l=|m|}^{l_{\max}} f_l^m Y_l^m(\theta(\mathbf{x}),0), \text{ and} \tag{2.1.15}$$

$$G_{m,y} = w(\mathbf{x}) \sum_{x=0}^{N_\theta-1} f(\mathbf{x}) \exp\left(-im\phi_{0,y} - \frac{2\pi imx}{N_{\phi,y}}\right), \tag{2.1.16}$$

the transforms then become

$$f(\mathbf{x}) = \sum_{m=-l_{\max}}^{l_{\max}} F_{m,y} \exp\left(im\phi_{0,y} + \frac{2\pi imx}{N_{\phi,y}}\right), \text{ and} \tag{2.1.17}$$

$$f_l^m = \sum_{y=0}^{N_\theta-1} G_{m,y} Y_l^m(\theta(\mathbf{x}),0). \tag{2.1.18}$$

Equations 2.1.16 and 2.1.17 can be calculated via a fast Fourier transform (FFT), so have worst case time complexity $O(N_\theta N_{\phi,y} \log(N_{\phi,y}))$ for prime $N_{\phi,y}$, however when $N_{\phi,y}$ prime factor decomposition consists only of small primes this can be improved. Therefore, the

time complexity of both the forward and backward transforms reduces to that of equations (2.1.15) and (2.1.18), namely $O(N_\theta l_{\max}^2)$, a factor of $\sim l_{\max}$ faster than brute force. For typical resolutions examined in this thesis the speedup will be around a factor of 200-300. Several libraries exist that efficiently implement these transforms and we will make use of that detailed in Reinecke & Seljebotn (2013).

It is worth noting that there exists an $O(l_{\max}^2 \log^2 l_{\max})$ implementation of the transforms for a pixel grid spaced evenly in θ and ϕ (Driscoll & Healy, 1994), but with a large prefactor on the computational complexity. This large prefactor makes the method desirable only when $l_{\max} \gtrsim 800$ so will not be used here.

2.1.3 Convolution on the sphere

Convolution of two functions f_1 and f_2 in d-dimensional, real coordinate space is uniquely defined as

$$(f_1 * f_2)(\mathbf{x}) = \int_{\mathbb{R}^d} f_1(\mathbf{y}) f_2(\mathbf{y}-\mathbf{x}) d\mathbf{y}, \tag{2.1.19}$$

with $\mathbf{x}$ and $\mathbf{y}$ denoting position vectors in this space.

By analogy, convolution on the 2-sphere can be defined (Wandelt & Górski, 2001) as

$$(f_1 * f_2)(g) = \int_{\mathbb{S}^2} f_1(\omega) \left[\Lambda(g) f_2\right](\omega) d\omega, \tag{2.1.20}$$

where g is a rotation in the rotation group SO(3), $\Lambda(g)$ are the rotation operators that rotate the functions through a rotation g and the integral is over all solid angles.

Performing spherical convolution numerically requires discretizing both the functions f_1 and f_2 and the rotations $g \in SO(3)$. With the functions represented in N_{pix} pixels and the rotations discretized into $l_{\max}$ rotations about each of the three Euler angles, this, brute force, approach to convolution would require $O(N_{\text{pix}} l_{\max}^3) \approx O(l_{\max}^5)$ operations. In the following two sections more efficient methods of calculating convolution are described.

Fast convolution of bandlimited functions

For the convolution of two functions, at least one of which is bandlimited, Wandelt & Górski (2001) describe an implementation that results in a speed up by a factor of

$O(\sqrt{N_{\text{pix}}})$ over the brute force method. Their algorithm involves factorizing the rotations in the convolution as follows. The spherical convolution in equation (2.1.20) is parameterized using the Euler angles as

$$(f_1 * f_2)(\alpha,\beta,\gamma) = \int_{\mathbb{S}^2} f_1(\omega)\left[\Lambda(\alpha,\beta,\gamma) f_2\right](\omega) d\omega \tag{2.1.21}$$

where α, β and γ are the size of the rotations of the object by the x, y and z axes, respectively. The rotation operator is factorized as

$$\Lambda(\alpha,\beta,\gamma) \equiv \Lambda\left(\alpha - \frac{\pi}{2}, \frac{-\pi}{2}, \gamma\right) \Lambda\left(0, \frac{\pi}{2}, \gamma + \frac{\pi}{2}\right), \tag{2.1.22}$$

which allows equation (2.1.24) to be written

$$(f_1 * f_2)(\alpha,\beta,\gamma) = \int_{\mathbb{S}^2} f_1(\omega) \left[\Lambda\left(\alpha - \frac{\pi}{2}, \frac{-\pi}{2}, \gamma\right) \Lambda\left(0, \frac{\pi}{2}, \gamma + \frac{\pi}{2}\right) f_2\right](\omega) d\omega \tag{2.1.23}$$

or substituting the spherical harmonic expansion of the functions f_1 and f_2 and using the orthogonality of the spherical harmonics given by equation (2.1.7), this can be written in spectral form as

$$(f_1 * f_2)(\alpha,\beta,\gamma) = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} (f_1)_l^m \left[\Lambda\left(\alpha - \frac{\pi}{2}, \frac{-\pi}{2}, \gamma\right) \Lambda\left(0, \frac{\pi}{2}, \gamma + \frac{\pi}{2}\right) f_2\right]_l^{m'}. \tag{2.1.24}$$

The rotation operators can be written, in spectral form, using the Wigner d-matrices $d^l_{mm'}$ as

$$\Lambda(\alpha,\beta,\gamma)^l_{mm'} = e^{-im\alpha} d^l_{mm'}(\beta) e^{-im'\gamma}. \tag{2.1.25}$$

These d-matrices are usually calculated by means of a recursion relation, such as that described in Risbo (1996). Using this decomposition, the convolution can be written

$$(f_1 * f_2)(\alpha,\beta,\gamma) = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l} \sum_{m'=-l}^{l} \sum_{m''=-l}^{l} d^l_{m'm}\left(\frac{\pi}{2}\right) d^l_{m'm''}\left(\frac{\pi}{2}\right) [f_1]^l_m [f_2]^l_{m''} \exp\left[im\left(\alpha - \frac{\pi}{2}\right) + im'\beta + im''\left(\gamma + \frac{\pi}{2}\right)\right]. \tag{2.1.26}$$

This calculation can be seen to require $O(l_{\max}^4)$ operations, an improvement by a factor of $l_{\max}$ over the method in the previous section.

Faster convolution with an azimuthally symmetric bandlimited function

The convolution between a circularly symmetric (or zonal) function (i.e. one that depends only on arc-length from a particular point, taken to be the north pole), G, with an arbitrary function, f, defined on the sphere can be written:

$$(G*f)(\boldsymbol{\xi}) = \int_{\mathbb{S}^2} G(\boldsymbol{\xi}\cdot\boldsymbol{\xi}')f(\boldsymbol{\xi}')d\boldsymbol{\xi}' \tag{2.1.27}$$

i.e. an integral over the 2-sphere.

The spherical harmonic (SH) representation of G is

$$G(\boldsymbol{\xi}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} g_l^m Y_l^m(\boldsymbol{\xi}). \tag{2.1.28}$$

As G is circular, $a_{lm} = 0$ for $m \neq 0$ thus

$$\begin{aligned} G(\boldsymbol{\xi}) &= \sum_{l=0}^{\infty} g_l^0 Y_0^l(\boldsymbol{\xi}) \\ G(\boldsymbol{\xi}) &= \sum_{l=0}^{\infty} \sqrt{\frac{2l+1}{4\pi}} g_l^0 P_l(\boldsymbol{\xi}), \end{aligned} \tag{2.1.29}$$

where P_l is the lth Legendre function. Substituting this back into equation (2.1.27) gives

$$\begin{aligned} (G*f)(\boldsymbol{\xi}) &= \sum_{l=0}^{\infty}\int_{S^2}\sqrt{\frac{2l+1}{4\pi}} g_l^0 P_l(\boldsymbol{\xi}\cdot\boldsymbol{\xi}')f(\boldsymbol{\xi}')d\boldsymbol{\xi}' \\ &= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sqrt{\frac{4\pi}{2l+1}} g_l^0 Y_l^m(\boldsymbol{\xi})\int_{S^2} Y_l^m(\boldsymbol{\xi}')f(\boldsymbol{\xi}')d\boldsymbol{\xi}' \\ &= \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sqrt{\frac{4\pi}{2l+1}} g_l^0 f_l^m Y_l^m(\boldsymbol{\xi}), \end{aligned} \tag{2.1.30}$$

which can be seen to be a simple scaling of the SH coefficients of the function f.

In this thesis we will often want to convolve with Gaussian functions. There are several ways of defining the a Gaussian-like function on the sphere, here we take the

fundamental solution of the heat equation to be our Gaussian. The heat equation is

$$\frac{\partial g}{\partial t} - k\nabla^2 g = 0, \tag{2.1.31}$$

where k is a positive constant. In 2-dimensional real coordinate space the fundamental solution (i.e. that with initial conditions of a single point source) to this equation is

$$g(\mathbf{x},t) = \frac{1}{\sqrt{4\pi kt}} \exp\left(-\frac{\mathbf{x}\cdot\mathbf{x}}{4kt}\right), \tag{2.1.32}$$

which is an 2-dimensional standard normal distribution with width $\sigma = \sqrt{2kt}$. Solving the heat equation on the sphere we find that the appropriate spherical harmonic coefficients of the solution are

$$g_l^m = \sqrt{\frac{2l+1}{4\pi}} e^{\frac{-l(l+1)\sigma^2}{2}}. \tag{2.1.33}$$

This algorithm has a time complexity of $O(l_{\max}^2)$ so in practice will be dominated by the time taken to perform the spherical harmonic transforms of the function being convolved, which in our implementation has $O(l_{\max}^3)$ complexity. This method is therefore $O(l_{\max})$ faster than the 'fast' method described in the previous section and $O(l_{\max}^2)$ faster than the brute force method. As, in this work, $l_{\max}$ will typically be $O(10^2)$, care will be taken to formulate the deconvolution techniques such that all convolutions are with azimuthally symmetric functions.

2.2 The problem with data

The process of observation can, in many cases, be described mathematically by stating that the data obtained, D, are the result of a convolution of the true, underlying, image, I, with the point spread function (PSF), B, of the detector along with the addition of some noise, N, i.e.

$$D(\mathbf{x}) = I * B(\mathbf{x}) + N(\mathbf{x}). \tag{2.2.34}$$

where the argument $\mathbf{x}$ refers to position or pixel number (as appropriate) and $*$ is the spatial convolution operator.

Along with this data we typically have, or can obtain, knowledge of the instrumental

PSF and noise. However, the quantity of interest is the true image, so we must then solve the inverse problem. Finding the best reconstruction, i.e. that closest to the original image, I, is the problem of image reconstruction. As the noise is known only statistically, equation (2.2.34) cannot be inverted analytically so we must instead adopt some statistical or iterative technique to perform the inversion.

Given that the number of degrees of freedom that may be involved in the problem is as large as the number of pixels that make up the image, the parameter space is very large and searching significant regions is computationally expensive. Several techniques have been developed, and popularized, to combat this problem including direct inversion using iterative procedures such as Jansson's method (discussed in the next section). However these techniques often lead to the production of spurious features in the reconstruction, driven by, and amplifying, noise in the data. To avoid these problems Bayesian image reconstruction techniques have been developed that make use of prior information, favouring simplicity, in addition to the measured data. These techniques are discussed in section 2.4.

We illustrate several of the methods described in this chapter using data based on NASA's 'blue marble' image[1]. The image consists of a set of true colour satellite images, taken between June and September, on land and shallow water, that are mosaiced across the globe. The image is defined on a sphere and consists of 350×175 pixels in an equirectangular grid. The oceans, including at the north pole are coloured dark blue. We created the data by blurring and adding Gaussian noise to each of the RGB colour channels in the image forming three 350×175 pixel arrays. The angular width, σ, of the blurring function is chosen to match that of that data on Mars used in later chapters. These reconstructions will be performed on each channel before recombining to make an image. The original image and blurred, noisy data are shown in figure 2.1. All of the 'blue marble' plots in this section will show as stereographic projections, but the data are defined, and the calculations performed, on the sphere.

[1]Credit and jpg available at http://visibleearth.nasa.gov/view.php?id=57752

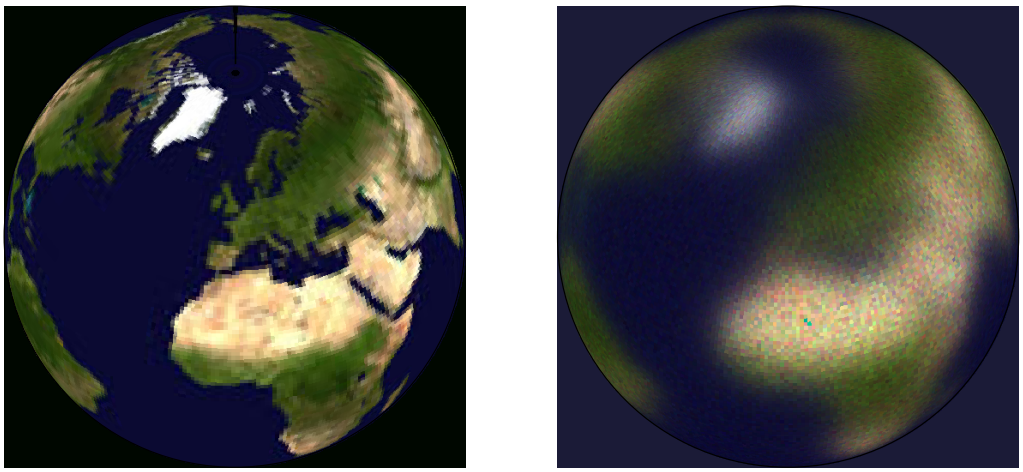

Figure 2.1: *Left:* Stereographic projection of the 'blue marble'. *Right:* 'Blue marble' blurred and with added noise.

2.3 Direct inversion and Van Cittert deconvolution

If one considers simply the problem of deconvolution perhaps the most obvious solution is to find the inverse of the PSF and convolve this with the data. In flat space, we can use the convolution theorem to write, in the absence of noise,

$$D(\mathbf{x}) = I * B(\mathbf{x}) \tag{2.3.35}$$

$$= \mathcal{F}^{-1}\{\tilde{I}\tilde{B}\} \tag{2.3.36}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform and the tilde indicates the Fourier transform of the object. For azimuthally symmetric PSFs the spherical analoge of the convolution theorem is given by equation (2.1.30). I can be seen to be

$$I(\mathbf{x}) = \mathcal{F}^{-1}\left\{\frac{\tilde{D}}{\tilde{B}}\right\}, \tag{2.3.37}$$

which in the presence of noise becomes

$$I(\mathbf{x}) = \mathcal{F}^{-1}\left\{\frac{\tilde{D}-\tilde{N}}{\tilde{B}}\right\}. \tag{2.3.38}$$

The PSF will often be bandlimited so its Fourier transform will be vanishingly small, thus some PSFs will not be invertible as their inverse will be infinitely large. This fact

along with the presence of noise can give rise to reconstructions that are dominated by large, spurious and often oscillatory features. A series of linear, iterative methods have been developed with the aim of controlling this undesired behaviour as it evolves. A prototypical example is Van Cittert deconvolution (van Cittert, 1931), which is defined by the equations

$$\hat{I}^{(0)}(\mathbf{x}) = D(\mathbf{x}), \tag{2.3.39}$$

$$\hat{I}^{(k+1)}(\mathbf{x}) = \hat{I}^{(k)}(\mathbf{x}) + (D(\mathbf{x}) - \hat{I}^{(k)} * B(\mathbf{x})). \tag{2.3.40}$$

That is, the new guess at the truth, $\hat{I}^{(k+1)}$, is the sum of the old guess, $\hat{I}^{(k)}$, and the residuals (corresponding to the kth guess) and the initial guess, $\hat{I}^{(0)}$, is set equal to the data. This method can be shown to be equivalent to direct Fourier inversion, truncated at kth order in the expansion of the function B. This truncation is, however, not sufficient to prevent the emergence of spurious features (Jansson, 1984).

2.3.1 Jansson's method

Jansson's method is similar to Van Cittert deconvolution but with an additional relaxation factor, r, modifying each iteration, the inclusion of this factor was motivated by the desire to suppress spurious noise features present in simple linear deconvolution techniques. The method is defined by the relations

$$\hat{I}^{(0)}(\mathbf{x}) = D * B(\mathbf{x}), \tag{2.3.41}$$

$$\hat{I}^{(k+1)}(\mathbf{x}) = I^{(k)}(\mathbf{x}) + r[I^{(k)}](D(\mathbf{x}) - I^{(k)} * B(\mathbf{x})), \tag{2.3.42}$$

$$r[I^{(k)}(\mathbf{x})] = r_0 \left(1 - 2 \left| \frac{I^{(k)}(\mathbf{x}) - I_{\min}}{I_{\max} - I_{\min}} - \frac{1}{2} \right| \right), \tag{2.3.43}$$

where r_0 is a relaxation constant whose value is determined empirically and $I_{\min}$ and $I_{\max}$ are the smallest and largest permitted values of the reconstruction. The form of r prevents solutions that are moving out of the allowed range of values from changing further as it goes to zero in pixels where the solution approaches $I_{\min}$ or $I_{\max}$. It was hoped that avoiding non-physical solutions (in the form of negative fluxes) would improve the final

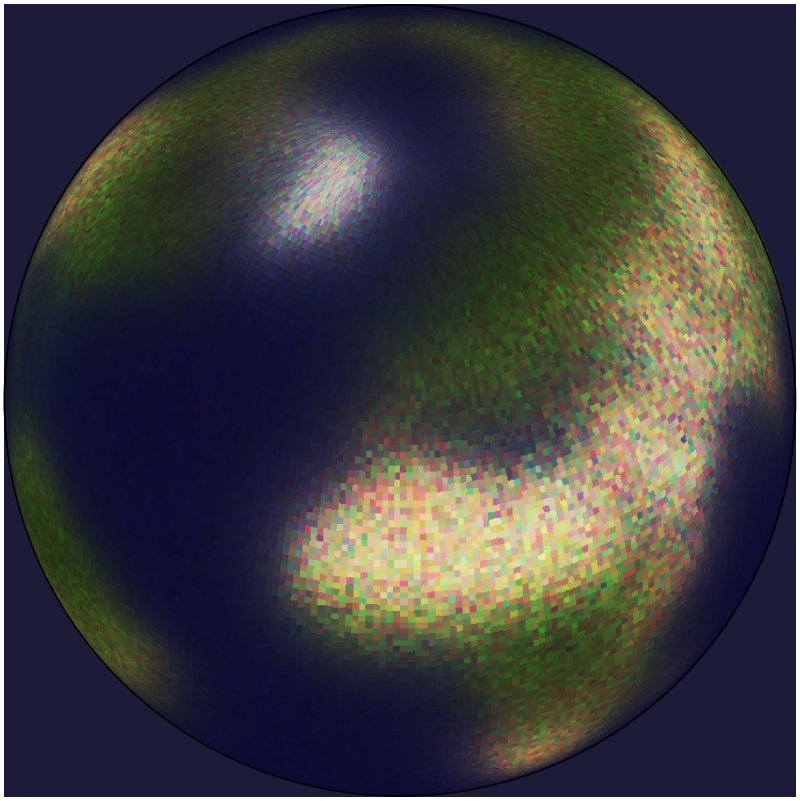

Figure 2.2: The result of a Jansson's method reconstruction of the data shown in figure 2.1.

image, even in regions outside of those truncated (Jansson, 1984). This method has seen some use in planetary sciences in the spatial deconvolution of neutron and gamma ray data sets (Elphic et al., 2005a,b; Lawrence et al., 2007; Prettyman et al., 2009). However it is still liable to produce spurious, noise-derived, features in the deconvolved images (Lawrence et al., 2007).

As repeated iterations of Jansson's method cause noise amplification, an appropriate maximum number of iterations must be found for each data set. In practice this is done by creating a mock set of data with statistics and, ideally, features that closely match the actual data. The method is then applied to this mock data and the number of iterations required to find the reconstruction most closely matching the known truth is determined. The method is then run on the actual data for the number of iterations found to be the most appropriate for the mock data.

The Jansson's method reconstruction of the test image is shown in figure 2.2. This image was created by stopping Jansson's method when the reconstruction was closest to the known truth (measured as a RMS difference between the RGB values in all pixels).

It therefore shows an ideal run, when applied to measured data the number of iterations will not be ideal so the image would appear smoother or noisier than presented here. Comparison with figure 2.1 shows some improvement in the range of colors reconstructed and sharpening at coastlines (most obvious around Greenland, i.e. the white area near the north pole). However, the reconstruction still greatly resembles the data, this is due to the amplification of noise leading to sharper reconstructions being disfavoured. To reconstruct the test image, Jansson's method took 37 seconds. All times in theis chapter are from runs using a single core of a 2.67 GHz Intel X5650.

2.4 Bayesian image reconstruction

An alternative approach to avoiding the emergence of spurious features is the use of Bayesian image reconstruction. In the methods described in the previous section, the aim was deconvolution and the suppression of noise was a secondary consideration, added to the algorithms to prevent the emergence of spurious features. Bayesian image reconstruction techniques start by asking a different question, one that is perhaps more relevant: given the data and our prior knowledge what is the most likely true, underlying image?

To answer this question we must consider how best to maximize the posterior conditional probability:

$$P(\hat{I},M|D) = \frac{P(D|\hat{I},M)P(\hat{I}|M)P(M)}{P(D)}, \tag{2.4.44}$$

where $\hat{I}$ is the inferred truth and M is the model, which describes the relationship between $\hat{I}$ and the data, including the PSF and the basis in which the image is represented. As the data are already taken, $P(D)$ is not affected by anything we can do so is constant. Additionally, to avoid bias, $P(M)$ will be assumed to be uniform. This assumption leaves two terms: the first, $P(D|\hat{I},M)$, is the likelihood of the data given a particular inferred truth and model, calculated using a goodness-of-fit statistic, for example χ^2 for data with Gaussian errors, where

$$P(D|\hat{I},M) = \exp(-0.5\chi^2), \tag{2.4.45}$$

and

$$\chi^2 = \sum \left(\frac{D(\mathbf{x}) - I * B(\mathbf{x})}{\sigma(\mathbf{x})} \right)^2, \tag{2.4.46}$$

where $\sigma(\mathbf{x})$ is the expected noise in pixel $\mathbf{x}$. For non-Gaussian errors the prior can be calculated directly from the probability density function. The second term, $P(\hat{I}|M)$, is the image prior. Simple maximum likelihood methods take this prior distribution to be uniform. However, given equation (2.2.34) (and enforcing the sum and product rules of probability) the prior can be shown to be (Garrett, 2001)

$$P(\hat{I}|M) \propto e^{\alpha S}, \tag{2.4.47}$$

$$S = \sum_i \left(c_i - m_i - c_i \ln\left(\frac{c_i}{m_i}\right)\right), \tag{2.4.48}$$

where α is a parameter setting the relative weight assigned to the prior against the data, S is the image entropy, the sum is over all independent image elements, c_i is the value of the image in element i and m_i is the model value, i.e. expected value given our prior knowledge. In the case of images with independent pixels the sum would be over the pixels, $\mathbf{x}$, with c_i equivalent to $\hat{I}(\mathbf{x})$.

Therefore for a set of reconstructed images with equal likelihoods the best reconstruction, i.e. that which maximizes the posterior, equation (2.4.44), will be that with maximum entropy. This maximum entropy criterion gives rise to a class of popular image reconstruction techniques and can be seen to follow naturally from a Bayesian interpretation of the problem of image reconstruction. The principle of maximum entropy can also be used to set the values of m_i (assuming that they must sum to some known value) by noting that the entropy is maximized when all m_i are equal.

2.4.1 Wiener deconvolution

Wiener deconvolution is an extension of the Wiener filter, which aims only to remove noise, to the problem of deconvolution (Bracewell, 1958; Helstrom, 1967). Although the original motivation for this method was to minimize the mean-square error in the reconstruction, Wiener deconvolution turns out to be the optimal Bayesian reconstruction when both the signal and noise are Gaussian random variables. Presented below is the Bayesian derivation of the filter (see e.g. Bracewell (1958) for a derivation based on minimizing errors). This derivation is simplified using a matrix representation of equation (2.2.34),

i.e.

$$\mathbf{d} = \mathbf{B}\mathbf{i} + \mathbf{n}, \tag{2.4.49}$$

where $\mathbf{d}$, $\mathbf{i}$ and $\mathbf{n}$ are 1-dimensional vectors containing the pixel values of the data, true image and noise, respectively, and $\mathbf{B}$ is a matrix representation of the operation of convolving with the PSF.

For an image made up of pixels with values given by a Gaussian random variable the image prior can be written

$$P(\hat{I}|M) \propto \exp(-\frac{1}{2}\hat{\mathbf{i}}\boldsymbol{\Sigma}^{-1}\hat{\mathbf{i}}), \tag{2.4.50}$$

where $\boldsymbol{\Sigma}$ is the a priori signal cross-correlation matrix i.e. $\boldsymbol{\Sigma} = \langle \mathbf{i}\mathbf{i}^{\intercal} \rangle$ and angled brackets represent an ensemble average. Therefore the posterior can be written

$$P(\hat{I}, M|D) \propto \exp(-\frac{1}{2}\chi^2 - \frac{1}{2}\hat{\mathbf{i}}\boldsymbol{\Sigma}^{-1}\hat{\mathbf{i}}), \tag{2.4.51}$$

which is maximized by minimizing $\chi^2 + \hat{\mathbf{i}}\boldsymbol{\Sigma}^{-1}\hat{\mathbf{i}}$. Noting that $\chi^2 = \mathbf{r}^{\intercal}\mathbf{N}^{-1}\mathbf{r}$, where $\mathbf{r} = \mathbf{d} - \mathbf{B}\hat{\mathbf{i}}$ and N the noise covariance matrix, i.e. $\mathbf{N} = \langle \mathbf{r}\mathbf{r}^{\intercal} \rangle$, and completing the square in $\hat{\mathbf{i}}$ we can write the posterior as

$$P(\hat{I}, M|D) \propto \exp(-\frac{1}{2}(\hat{\mathbf{i}} - \bar{\mathbf{i}})^{\intercal}\boldsymbol{E}^{-1}(\hat{\mathbf{i}} - \bar{\mathbf{i}})), \tag{2.4.52}$$

with

$$\bar{\mathbf{i}} = \mathbf{W}\mathbf{d}, \tag{2.4.53}$$

$$\mathbf{E} = (\boldsymbol{\Sigma}^{-1} + \mathbf{r}^{\intercal}\mathbf{N}^{-1}\mathbf{r})^{-1}, \tag{2.4.54}$$

$$\mathbf{W} = \mathbf{E}\mathbf{r}^{\intercal}\mathbf{N}^{-1}. \tag{2.4.55}$$

Clearly, the reconstruction that maximizes this posterior is $\hat{\mathbf{i}} = \bar{\mathbf{i}}$. It is often more practical to perform this filtering in the spectral domain where the filter is given by (Zaroubi et al., 1995)

$$\tilde{W} = \frac{\tilde{B}^* C_I}{|\tilde{B}|^2 C_I + C_N}, \tag{2.4.56}$$

where the tilde indicates the spectral domain (i.e. Fourier or spherical harmonic trans-

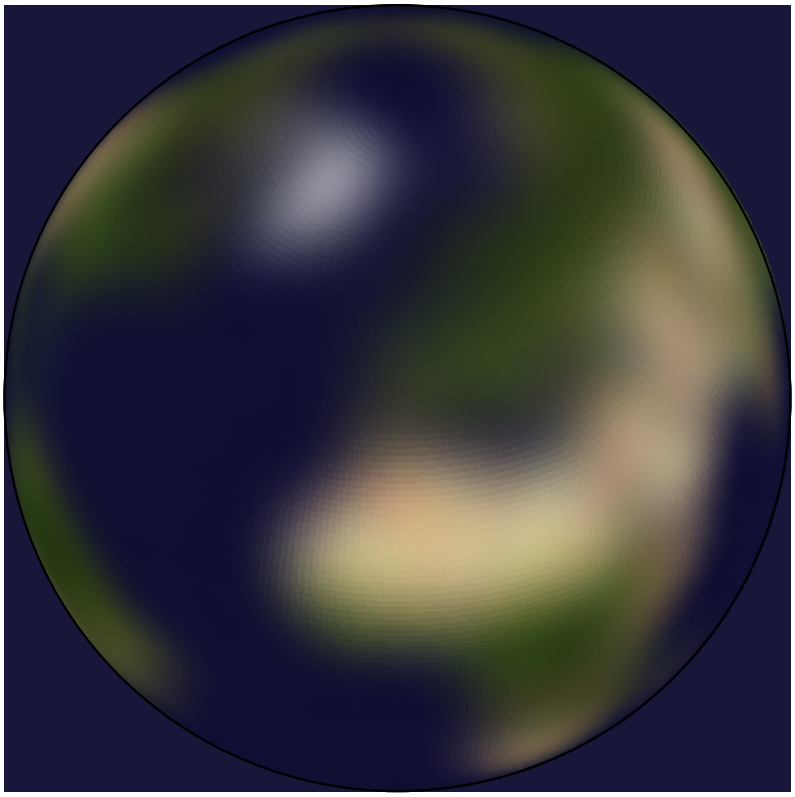

Figure 2.3: The result of a Wiener deconvolution of the data shown in figure 2.1.

form) and C_I and C_N are the power spectra of the signal and noise, respectively.

As Weiner deconvolution consists of a single convolution it is the fastest of the methods discussed here. The reconstruction of the test image, shown in figure 2.3, taking less than one second to complete. However, for the Wiener deconvolution technique to be effective the covariance, or power spectra, of the signal and noise must be known before the deconvolution is performed, which will often not be the case. The reconstruction shown in figure 2.3 was performed in the spectral domain and the noise was assumed to be white and the signal to have a power law power spectrum. The method effectively suppresses the noise, but the power law prior has the effect of smoothing the reconstruction.

2.5 Pixon image reconstruction

Many Bayesian image reconstruction techniques assume that the model, M, in equation (2.4.44) is fixed. However, this is not a necessary assumption, allowing both I and M to vary in the reconstruction enables the prior to be interpreted as a mathematical state-

ment of Occam's razor. More complex models, allowing a larger number of reconstructed truths, are naturally disfavoured and allowed only if they provide a sufficiently good improvement in the likelihood. In the pixon method, the model, that is how the pixels are arranged, is allowed to vary. It is an adaptive image reconstruction technique, in which the reconstructed image is described on a grid of pixons, where a pixon is a collection of pixels whose shape and size are allowed to vary. Thus areas of the image containing little information are described by a few large pixons, whereas regions of the data containing more information are described by smaller pixons, giving the reconstruction the freedom to vary on smaller scales.

This method is motivated by considering an image as a collection of c indistinguishable events in n cells. What these events and cells are is a choice to be made when devising a reconstruction technique, though natural choices include the number of observations, total signal to noise ratio or reconstructed value within each pixel of the reconstructed image. The probability of placing c_1 events in the first cell is given by the binomial distribution:

$$P(c_1|c,n) = \frac{c!}{c_1!(c-c_1!)} p^{c_1}(1-p)^{c-c_1}. \tag{2.5.57}$$

Assigning an equal probability of placing an event in any cell, $p = \frac{1}{n}$, gives

$$P(c_1|c,n) = \frac{c!}{c_1!(c-c_1!)} \frac{(n-1)^{c-c_1}}{n^c}. \tag{2.5.58}$$

The probability of finding c_2 events in the second cell is then

$$P(c_2|c,n,c_1) = \frac{(c-c_1)!}{c_2!(c-c_1-c_2!)} \frac{(n-2)^{c-c_1-c_2}}{(n-1)^{c-c_1}}. \tag{2.5.59}$$

The probability of a set of events in cells is therefore:

$$P(\{c_i\}|n) = \frac{c!}{n^c \prod_{i=0}^{n} c_i!}, \tag{2.5.60}$$

where c_i are the total number of events in cell i. Using Stirling's approximation, this can

be written as

$$\ln(P(\{c_i\}|n)) = \exp(\ln(c!) - c\ln(n) - \sum_i \ln(c_i!)), \tag{2.5.61}$$

$$\approx c\ln\left(\frac{c}{n}\right) - \sum_i c_i \ln(c_i), \tag{2.5.62}$$

$$\approx \sum_i c_i \ln\left(\frac{c}{n}\right) - c_i \ln(c_i) \tag{2.5.63}$$

$$\approx \sum_i c_i - \frac{c}{n} - c_i \ln\left(\frac{c_i n}{c}\right), \tag{2.5.64}$$

which is identical, with the substitution $m_i = \frac{c}{n}$, to equation (2.4.47).

It was noted by Pina & Puetter (1993) that each cell, or pixel, represents a degree of freedom and in regions of the image with low signal to noise ratio or containing little structure this may cause the image to be over specified. Grouping sets of pixels together would remove this problem and greatly improve the prior, which is strongly dependent on the number of cells. It is these groups of pixels, called pixons, that form the fundamental image units in pixon reconstructions. For an image made up of n pixons and containing C separate and indistinguishable detections, the probability of observing I_i detections in pixon i is

$$P(\hat{I}|M) = \frac{C!}{n^C \Pi_{i=1}^n I_i!}. \tag{2.5.65}$$

The problem of image reconstruction is then reduced to finding the pixon basis that maximizes the prior whilst also allowing an acceptable likelihood to be found. The prior is maximized, for a given number of pixons, by having the same information content in each pixon, i.e. $I_i = C/n, \forall i$. The image prior increases as fewer pixons are used, causing the image reconstruction to yield an $\hat{I}$ that contains the least possible structure whilst still being consistent with the data.

The reconstructed image that maximizes the posterior probability should provide both a good fit to the data and be sufficiently simple that overfitting of the data is prevented. In the pixon method, this overfitting is prevented by describing the image using a set of pixons instead of the pixels within which the data are defined. There will be far fewer pixons than pixels in the image, which results in a much simplified image and improved prior probability.

The pixon method (Pina & Puetter, 1993) has successfully been used in a range of disciplines including medical imaging, IR and X-ray astronomy [Puetter 1996 and references therein]. In addition, it has recently been used to reconstruct remotely sensed neutron (Eke et al., 2009) and gamma ray data (Lawrence et al., 2007) and has been shown to give a spatial resolution 1.5-2 times that of other methods in reconstructing planetary data sets (Lawrence et al., 2007).

All previous implementations of the pixon method avoid the computationally costly problem of assigning sets of pixels to distinct pixons by instead grouping pixels via convolution with a smoothing kernel, the size of which may vary across the image. The following two subsections will describe such methods. Firstly, the conventional 'Maximum entropy' implementation described in Pina & Puetter (1993) and Eke (2001) is detailed. Secondly, a new implementation in which the smoothing kernel sizes are set by considering a local misfit statistic is detailed. This second method will be called the 'locally adaptive pixon method'. Finally, in a third subsection, another new implementation of the pixon method is described, wherein the pixons are formed from collections of pixels without convolution with a smoothing kernel.

2.5.1 Maximum entropy pixon reconstruction

The maximum entropy pixon method aims to find the reconstruction that best maximizes the posterior probability, equation (2.4.44), by enforcing maximization of the image prior, equation (2.5.65), for a given number of pixons. Although this may not be optimal it greatly simplifies calculation as the size of the smoothing kernel, in each pixel, is determined only by the data and number of pixons.

The implementation of maximum entropy pixons used in this thesis is based on the speedy pixon method described in Eke (2001) and Eke et al. (2009). In this realization the reconstruction, $\hat{I}$, is based on a 'pseudoimage' that is defined in the same pixel grid as the data. $\hat{I}$ is constructed by convolving this pseudoimage, H, with a Gaussian kernel, K, whose width, δ, may vary across the image, i.e.

$$\hat{I}(\mathbf{x}) = (K_{\delta(\mathbf{x})} * H)(\mathbf{x}). \tag{2.5.66}$$

The local smoothing scale, or pixon width, $\delta(\mathbf{x})$, is determined for each pixel such that the information content is constant in each pixon over the entire image (given by $\Upsilon(\mathbf{x}) = (K_{\delta(\mathbf{x})} * \upsilon)(\mathbf{x})$, where $\upsilon(\mathbf{x})$ is the signal-to-noise ratio in pixel $\mathbf{x}$), i.e. the elements in equation (2.5.65) are units of signal to noise ratio. Setting the pixon sizes this way makes this form of pixon reconstitution a perfect maximum entropy method, that is the entropy is described in such a way that it is maximal by definition.

In practice, a finite set of distinct pixon sizes is used as this keeps the time of a single calculation of of the misfit statistic down to $O(n_{\text{pixels}} \log n_{\text{pixels}})$ (i.e. that of a fast Fourier transform), where n_{pixels} is the number of pixels in the image, whereas if the pixon size were allowed to vary continuously then the time would be $O(n_{\text{pixels}}^2 \log n_{\text{pixels}})$. In order to generate $\hat{I}$ we interpolate linearly between the images based on the two pixon sizes closest to those required for each pixel.

The global misfit statistic, $\mathbf{E_R}$

The misfit statistic is derived from the reduced residuals between the data and the blurred model,

$$R(\mathbf{x}) = \frac{D(\mathbf{x}) - (\hat{I} * B)(\mathbf{x})}{\sigma(\mathbf{x})}, \tag{2.5.67}$$

where $\sigma(\mathbf{x})$ is the anticipated statistical noise in pixel $\mathbf{x}$. Rather than using $\chi^2 = \sum R^2$, we adopt E_R from Pina & Puetter (1992) as the misfit statistic. This statistic is defined as

$$E_R = \sum_{\mathbf{y}=0}^{m} A_R(\mathbf{y})^2, \tag{2.5.68}$$

where A_R is the autocorrelation of the residuals, $A_R(\mathbf{y}) = \int R^*(\mathbf{x}) R(\mathbf{x}+\mathbf{y}) \mathrm{d}\mathbf{x})$ for a pixel separation, or lag, of $\mathbf{y}$. The benefit of minimizing E_R, over the more conventional χ^2, is that doing so suppresses spatial correlations in the residuals, preventing spurious features being formed by the reconstruction process. Pina & Puetter (1992) recommend that the autocorrelation terms defining E_R should be those corresponding to pixel separations smaller than the instrumental PSF. For a well-sampled PSF this means many different pixel lags. However, for the reconstructions we have attempted, there is negligible difference between those including different numbers of pixel lags. We therefore use only the four distinct terms with adjacent pixels (including diagonally adjacent pixels in 2D)

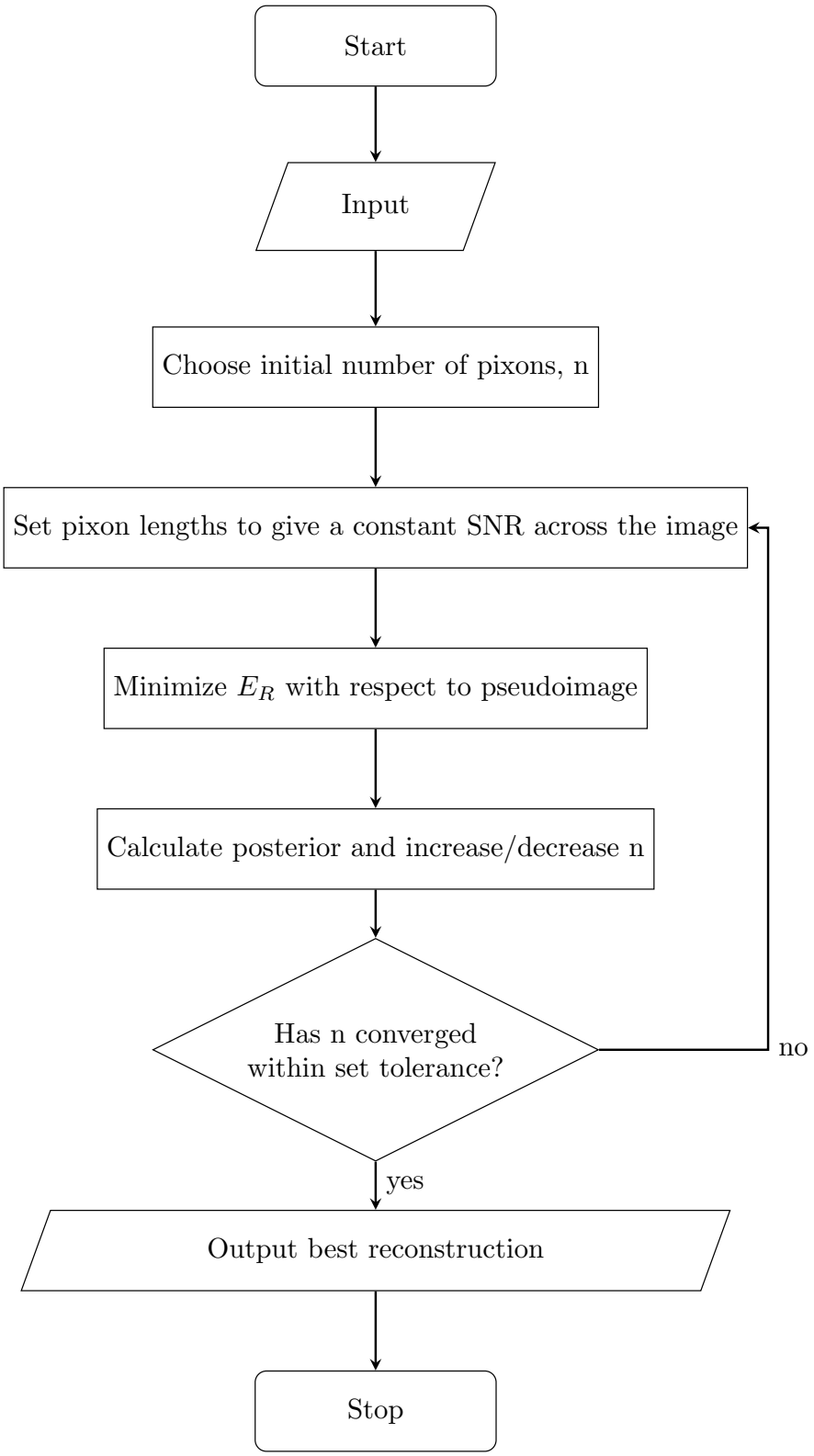

Figure 2.4: A flow chart illustrating the maximum entropy pixon image reconstruction algorithm.

to speed up the computation.

Prior to this work the pixon method has been implemented only on flat planes. When calculating E_R on a spherical surface we decided to map the sphere to a flat (i.e. Gaussian curvature equal to zero) torus. This is done via an equirectangular projection of the sphere to a square Cartesian plane (by extending the regular parameterization of the sphere to $0 < \theta < 2\pi$, $0 < \phi < 2\pi$) and identifying oposite edges. This mapping leaves the details of the calculation, including the use of FFTs for fast convolution, unchanged from the flat 2d case. However, as the lag terms included in E_R cover a certain number of pixels this mapping has the side effect of mixing the physical scales over which the statistic is defined.

Structure of the pixon algorithm

The maximization of the posterior (equation (2.4.44) is done iteratively, with two stages in each iteration (as illustrated in figure 2.4). Firstly, for a given number of pixons, the pixon sizes as a function of position are set so as to maximize the image prior (equation 2.5.65). Secondly, the values in the pseudoimage are adjusted to minimise the misfit statistic using the Polak-Ribière conjugate gradient minimization algorithm (Press et al., 1992).

In subsequent iterations, the number of pixons is varied using the golden section search method (Press et al., 1992) to maximize the posterior probability, or in practice its logarithm, which by combining equations (2.4.44-2.5.65) and using Stirling's approximation gives

$$\begin{aligned} \ln(P(\hat{I},M|D)) = \quad & \frac{1}{2}(\ln(N) + (1-n)\ln(2\pi) \\ & -n\ln(N/n) - \chi^2). \end{aligned} \tag{2.5.69}$$

Performance of the maximum entropy pixon algorithm

The result of performing a maximum entropy pixon reconstruction on the test image data is shown in figure 2.5. This reconstruction is clearly an improvement over Janssons's method and the Wiener filter. The Bayesian method is much more effective at removing noise than Jansson's method and provides a more detailed reconstruction than both previous techniques. The cost of this improvement is the greater computation complexity

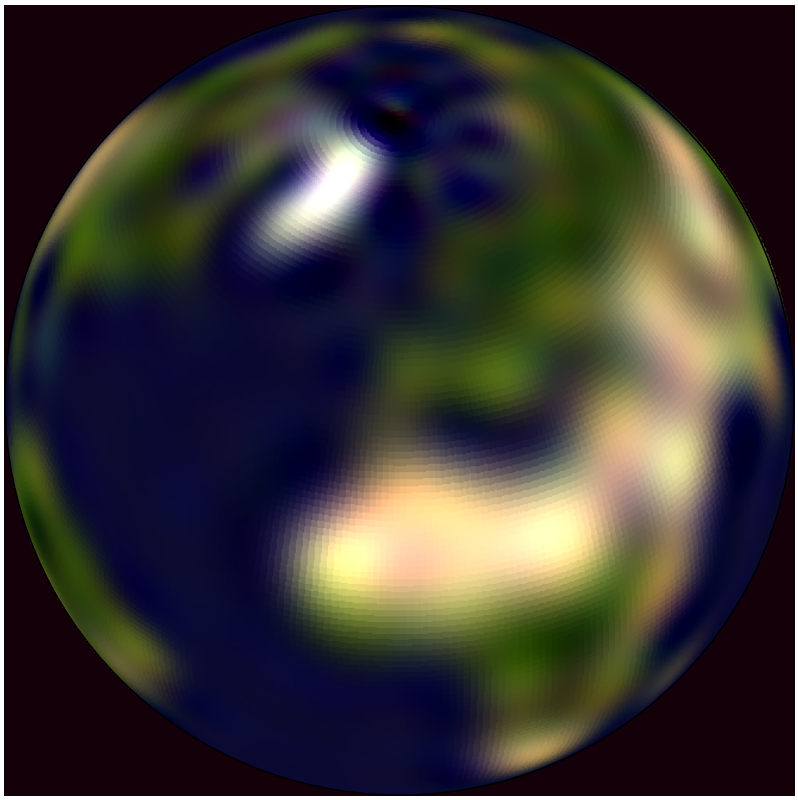

Figure 2.5: The result of a maximum entropy pixon reconstruction of the data shown in figure 2.1.

and therefore time required to perform the reconstruction. To produce the reconstruction in figure 2.5 took around six hours, two orders of magnitude more than required for the previous methods.

2.5.2 Locally adaptive pixon reconstruction

In this section the pixon method used in chapter 4 is described. Previous versions of the pixon method imposed a strict maximum entropy constraint on the solution (Pina & Puetter, 1993; Eke, 2001), in order to simplify the maximization of the posterior by reducing the parameter space that is to be explored. Here we make no such requirement, which will allow the creation of previously inaccessible and more likely reconstructions.

As in the maximum entropy pixon method the reconstructed count-rate map, $\hat{I}$, is based on a pseudoimage that is defined in the same grid as the reconstructed image, however this need not be the same as the data grid. This allows the reconstruction to be much more detailed than the data grid, which will usually be set only to sample well the

scale set by the PSF. As in the maximum entropy case, $\hat{I}$ is constructed by convolving the pseudoimage, H, with a Gaussian kernel, K, whose width, δ, may vary across the image, i.e.

$$\hat{I}(\mathbf{x}) = (K_{\delta(\mathbf{x})} * H)(\mathbf{x}). \tag{2.5.70}$$

Initially the smoothing scale, or pixon width, $\delta(\mathbf{x})$, is large. The final value is determined for each pixel using an iterative procedure, which includes the calculation of a newly defined local misfit statistic. Again, for reasons of speed, a distinct set of pixon widths is used and $\hat{I}$ is generated by interpolating linearly between the images based on the two pixon sizes closest to those required for each pixel.

Local misfit statistic, $\mathbf{E'_R}$

To create a local E_R statistic, E'_R, we begin by localizing A_R by multiplying with a kernel, K, such that

$$A'_R(\mathbf{x},\mathbf{z}) = \int R(\mathbf{y})R(\mathbf{x}+\mathbf{y})K(\mathbf{z}-\mathbf{y})\,d\mathbf{y}, \tag{2.5.71}$$

which we can rewrite in a form resembling a convolution, by defining $C_{\mathbf{x}}(\mathbf{y}) = R(\mathbf{y})R(\mathbf{x}+\mathbf{y})$, as

$$A'_R(\mathbf{x},\mathbf{z}) = \int C_{\mathbf{x}}(\mathbf{y})K(\mathbf{z}-\mathbf{y})\,d\mathbf{y}. \tag{2.5.72}$$

We are free to choose the kernel and, for ease of calculation, select a Gaussian, the width of which should be larger than the largest pixon width so that it is the pixon size, and not its gradient, that sets the smoothing scale in the reconstruction. For use in determining pixon widths we need to know the distribution of A'_R. Noting that $C_{\mathbf{x}}(\mathbf{y})$ has a normal product distribution with width one and applying the central limit theorem allows one to deduce that A'_R must have a standard normal distribution so E'_R obeys a χ^2 distribution.

When convolving in angular space we should define (using equation **??**).

$$A'_R(\mathbf{x},\omega(\mathbf{z})) = \left(\frac{C_{\mathbf{x}}}{a} * K\right)(\omega(\mathbf{z})), \tag{2.5.73}$$

and

$$E'_R(\mathbf{z}) = \sum_{\mathbf{y}} \frac{{A'_R}^2(\mathbf{y},\mathbf{z})}{\sigma^2_{A'_R}(\mathbf{z})}, \tag{2.5.74}$$

where $\sigma^2_{A'_R}$ can be deduced by analogy with a random walk i.e.

$$\begin{aligned} \sigma^2_{A'_R}(\mathbf{z}) &= \int K^2(\mathbf{z}-\mathbf{y})\, d\mathbf{y} \\ &= \left(\frac{1}{a} * K^2\right)(\omega(\mathbf{z})). \end{aligned} \tag{2.5.75}$$

Prior calculation

In the pixon method described in the previous section the content of each pixon was defined to be equal, so its explicit numerical calculation is never necessary. However, our new pixons are not only inhomogeneous but there needn't be an integer number (with pixon number defined as

$$n = \sum_{\mathbf{x}} \frac{1}{2\pi\delta(\mathbf{x})^2}, \tag{2.5.76}$$

i.e. the number of independent degrees of freedom in the image). For the case of a non-integer number of pixons we choose to weight the contribution of the partial pixon, such that the natural logarithm of the prior (using Stirling's approximation) becomes

$$\begin{aligned} \ln(P(I|M)) \approx{}& N\ln\left(\frac{N}{n}\right) + \frac{1}{2}\ln(N) - \frac{n-1}{2}\ln(2\pi) \\ &- \sum_{i=1}^{\lfloor n \rfloor}\left(N_i + \frac{1}{2}\right)\ln(N_i) - (n - \lfloor n \rfloor)\left(\left(N_n + \frac{1}{2}\right)\ln(N_n)\right), \end{aligned} \tag{2.5.77}$$

where N_n is the information in the final fraction of a pixon linearly scaled up to the size it would be if it were filling a whole pixon.

Structure of the locally adaptive pixon algorithm

Our goal is to maximize the posterior probability (equation 2.4.44) whilst ensuring that the residual field should be indistinguishable from a random Gaussian field. We do this using an iterative procedure with two stages in each iteration (figure 2.6). Firstly, for a given pixon size distribution the values in the pseudoimage are adjusted to minimize the global misfit statistic (section 2.5.1) using the Polak-Ribière conjugate gradient minimiza-

tion algorithm (Press et al., 1992). Secondly, we calculate the value of the posterior and, for each pixel in the data-grid, the local misfit statistic (section 2.5.2). These local misfit statistics are then ordered to form their cumulative distribution function (CDF). The difference, in each pixel, between the actual and expected, χ^2-distributed, CDF is then used to modify the pixon size in the same pixel, illustrated in figure 2.7. The change in pixon width is directly proportional to this difference, with the constant of proportionality determined empirically such that structure appears in the image sufficiently slowly to avoid introducing spurious features.

The combination of global misfit statistic minimization and pixon size modification is repeated, starting with large pixons and consequently a flat, constant reconstructed image everywhere, until the posterior probability stops increasing.

Performance of the locally adaptive pixon algorithm

The result of performing a locally adaptive pixon reconstruction on the test image data is shown in figure 2.8. As with the maximum entropy pixon method, the reconstruction supresses noise more effective than Jansson's method and produces a sharper image than Jansson's and Wiener decomposition. Producing the reconstruction in figure 2.8 took around 12 hours, around twice the time of the maximum entropy method. This increase in computation time is due to the extra work required to calculate the local misfit statistic and the larger number of iterations taken to find the final solution. The number of iterations required is, in part, dependent on the scaling factor used to convert the measured misfit statistic to a change in pixon width. A larger factor results in larger steps in pixon size and faster convergence but smaller factors may give a better result.

For this test data the two pixon methods produce very similar results, however the locally adaptive method produces less spurious structure in the oceans, which in the test image are a constant colour.

2.5.3 Hard pixons

Implementing the pixon problem as it was originally described, i.e. grouping discrete sets of pixels into pixons, is a computationally difficult task. Specifically, it belongs to the class of problems in combinatorial optimization, such as the travelling salesman problem,

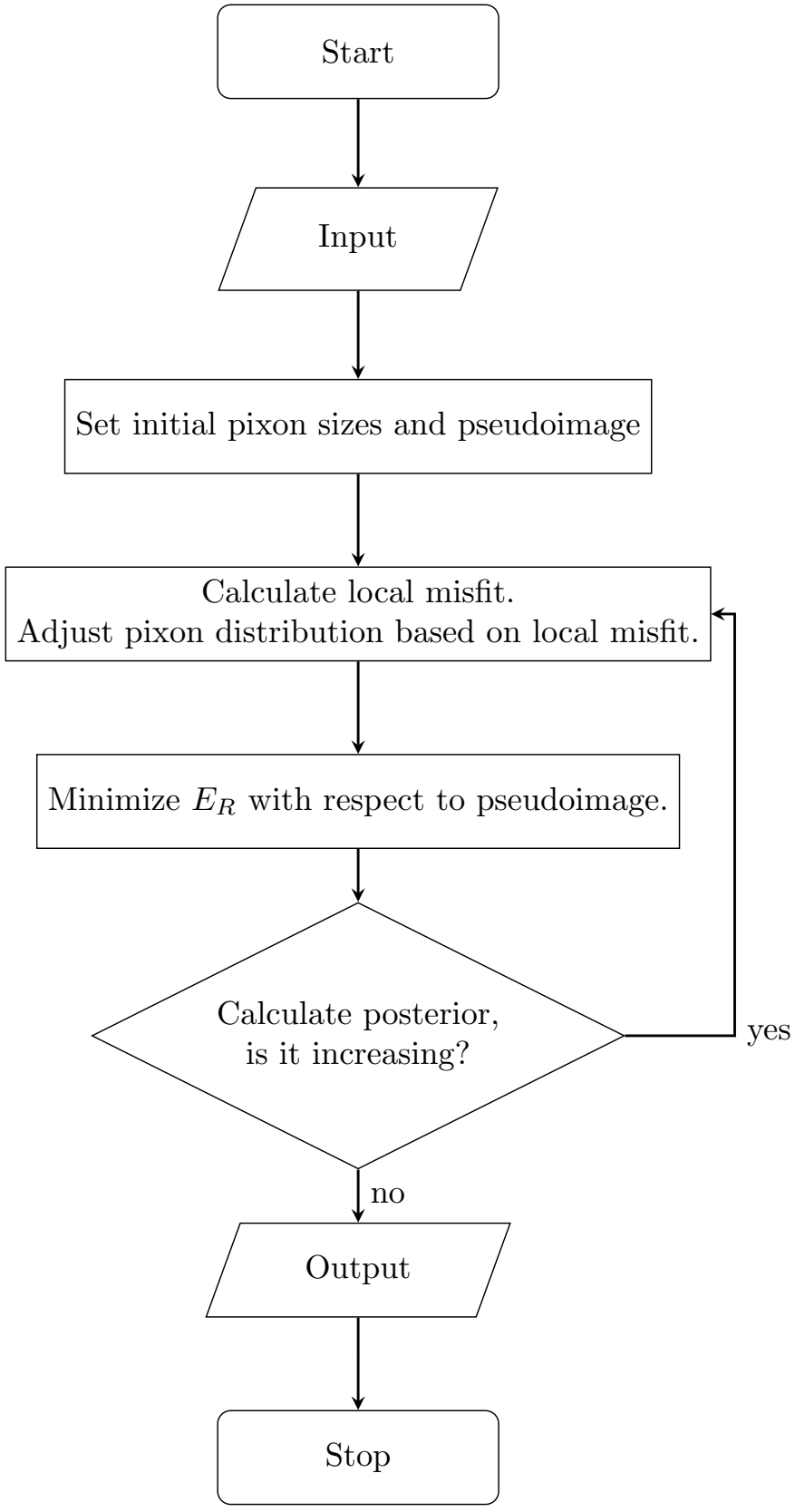

Figure 2.6: A flow chart showing the locally adaptive pixon image reconstruction algorithm.

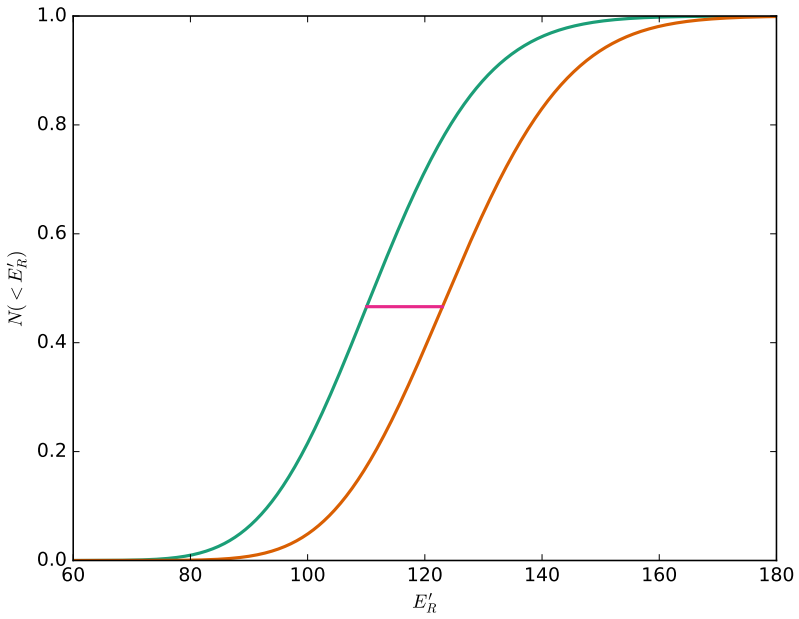

Figure 2.7: The expected, green, and observed, orange, cumulative distribution functions of the E'_R statistic. The size of the discrepancy between the two distributions, used to scale the change in pixon size for a particular pixel, is shown in red.

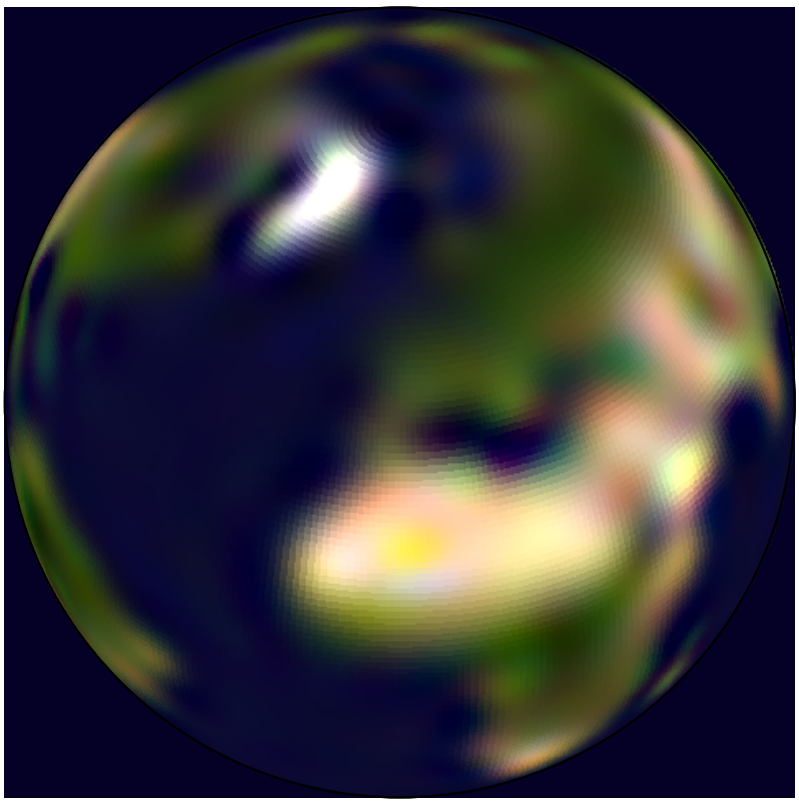

Figure 2.8: A locally adaptive pixon reconstruction of the data shown in figure 2.1.

whose worst case running time may increase non-polynomially. However, it is possible to find a good, if not strictly optimal solution by making some simplifying assumptions. An approach to the pixon problem, with simplifying assumptions, is outlined below.

Each pixel is assigned a pixon number, p, between zero and the number of pixons, n and each pixon a value c_i, i.e.

$$p(\mathbf{x}) = i \qquad 0 \leq i < n \tag{2.5.78}$$

$$c_i = c. \tag{2.5.79}$$

The pixons are kept ordered such that $c_i \leq c_j$ for $i < j$, which may require relabelling the pixons. This is done so that the gradients in the prior and misfit statistic may be calculated with respect to pixon number. The reconstructed image is then written

$$\hat{I}(\mathbf{x}) = c_{p(\mathbf{x})}, \tag{2.5.80}$$

and the residuals and misfit statistics have their usual definitions. The calculation of the prior is much simplified and can be shown by counting arguments to be given by equation (2.5.65). We can maximize the posterior, given by (equation (2.4.44)), with respect to $p(\mathbf{x})$ and c_i using conjugate gradient minimisation if the gradients of the prior and likelihood are known.

In the fuzzy pixon methods the gradient in the prior with respect to the pseudoimage and pixon sizes cannot be calculated as pixels are not assigned to particular pixons. However, with hard pixons this is not the case as each pixel really is associated with a pixon so we can calculate the change to the prior due to moving pixels between pixons and changing the value within a pixon. First we consider the log of the prior (equation (2.5.65)), expanded using Stirling's approximation

$$\ln(P(I|M)) \approx C \ln\left(\frac{C}{n}\right) + \frac{1}{2}\ln(C) - \frac{n-1}{2}\ln(2\pi) - \sum_{i=1}^{n}\left(c_i + \frac{1}{2}\right)\ln(c_i), \tag{2.5.81}$$

where $C = \sum_i c_i$. This prior is independent of $p(\mathbf{x})$ and differentiating with respect to c_j

gives

$$\frac{d\ln(P(I|M))}{dc_j} \approx \ln\left(\frac{C}{c_j n}\right) + \frac{1}{2}\left(\frac{1}{C} - \frac{1}{c_j}\right). \tag{2.5.82}$$

The gradient of the familiar misfit statistics can also be calculated. In the case of pixon value this is straightforward, note,

$$\frac{d\hat{I}(\mathbf{y})}{dc_j} = M_j(\mathbf{y}) \equiv \begin{cases} 1, & \text{if } p(\mathbf{y}) = j \\ 0, & \text{otherwise} \end{cases} \tag{2.5.83}$$

and

$$\frac{dR(\mathbf{y})}{dc_j} = -\frac{(M_j * B)(\mathbf{y})}{\sigma(\mathbf{y})}. \tag{2.5.84}$$

Therefore as

$$\chi^2 = \sum_{\mathbf{x}} R(\mathbf{x})^2, \tag{2.5.85}$$

its gradient is

$$\frac{d\chi^2}{dc_j} = -2\sum_{\mathbf{x}} R(\mathbf{x}) \frac{(M_j * B)(\mathbf{x})}{\sigma(\mathbf{x})}. \tag{2.5.86}$$

Calculating the gradient with respect to pixon number requires one to treat the variable as though it were continuous before rounding to the nearest integer at the end of the calculation. We can approximate the image gradient as

$$\frac{d\hat{I}(\mathbf{y})}{dp(\mathbf{x})} = \frac{dc_{p(\mathbf{x})}}{dp(\mathbf{x})}\delta_{\mathbf{xy}} \tag{2.5.87}$$

$$\approx (c_{p(\mathbf{x})+1} - c_{p(\mathbf{x})})\delta_{\mathbf{xy}} \tag{2.5.88}$$

$$\equiv \Delta(\mathbf{x})\delta_{\mathbf{xy}} \tag{2.5.89}$$

so

$$\frac{dR(\mathbf{y})}{dp(\mathbf{x})} = \Delta(\mathbf{x})B(\mathbf{x}-\mathbf{y}) \tag{2.5.90}$$

and

$$\frac{d\chi^2}{dp(\mathbf{x})} = -2\Delta(\mathbf{x})(R * B)(\mathbf{x}). \tag{2.5.91}$$

Equations (2.5.82), (2.5.86), and (2.5.91) can then be summed to find the gradient of the posterior with respect to all variables. The procedure for maximizing the posterior involves simply varying n to maximize the posterior, which is maximized for a given n

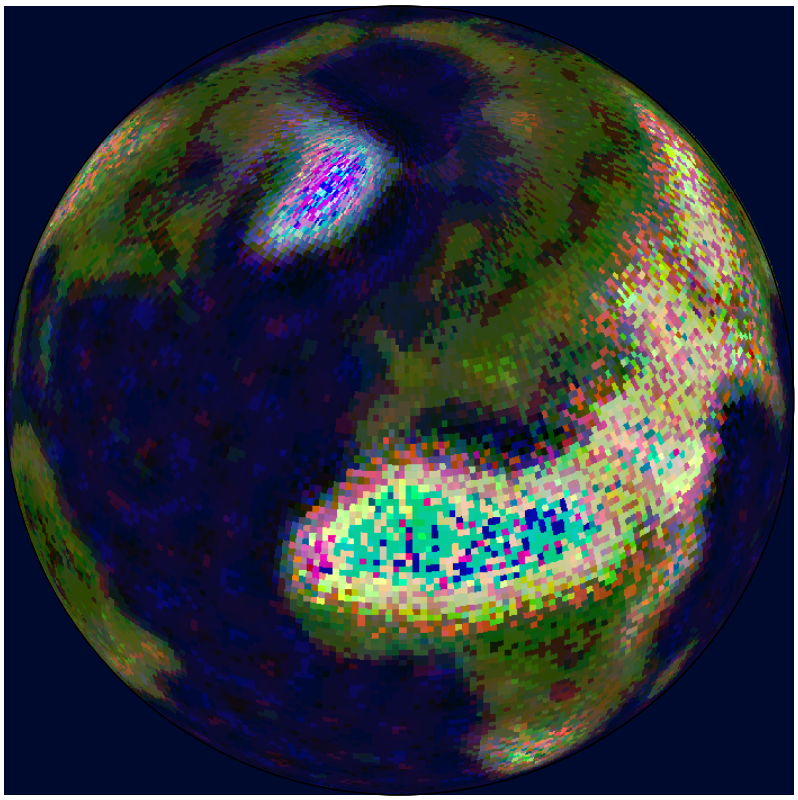

Figure 2.9: A hard edged pixon reconstruction of the data shown in figure 2.1.

using a conjugate gradient method.

The result of performing a hard pixon reconstruction on the test image data is shown in figure 2.9. Unlike the pixon methods in the previous section the effect of noise on this reconstruction is immediately obvious. However, behind the noise more structure is evident than before. Great Britain is seen distinctly (though not as an island) and Scandinavia is much clearer. The bright colours in the Sahara and Greenland are due to the color values going beyond the range allowed. This is the fastest of the pixon methods, taking around 40 minutes to complete its run.

2.6 Comparison of image reconstruction methods

To compare these reconstructions we define (by analogy with χ^2) the misfit statistic ψ^2,

$$\psi^2 = \sum_{\mathbf{x}} \frac{(I(\mathbf{x}) - \hat{I}(\mathbf{x}))^2}{\sigma(\mathbf{x})^2}, \tag{2.6.92}$$

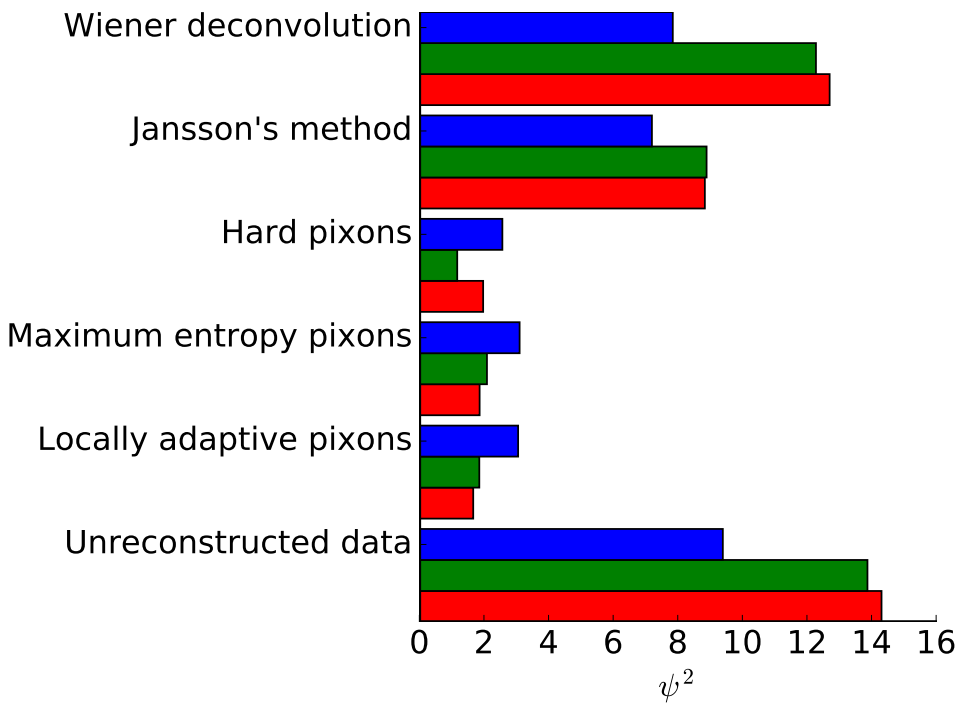

Figure 2.10: The ψ^2 misfit statistic calculated for several reconstructions of the test data using different techniques, labelled on the left. The red, green and blue bars correspond to reconstructions of the red, green and blue colour channels of the test image, respectively.

The values of this statistic for the different reconstructions are shown in figure 2.10. It is clear from this figure that the Bayesian, pixon methods produce reconstructions closer to the true image than the other methods examined. Although the difference is small, for this test data, the locally adaptive pixon reconstructions are better, in all colour channels, than those derived using the maximum entropy pixon method. Surprisingly, given the apparent noise in figure 2.9, the hard pixon method proves the most successful in reconstructing the test data. In this instance the ability of this method to create sharp boundaries, such as those at the coastlines, more than makes up for any noise that remains in the reconstructed image. These results are an example of the relative merits of the different reconstruction methods but the optimum reconstruction method in a particular situation will be determined by the data and the nature of the true image.

CHAPTER 3

Evidence for explosive silicic volcanism on the Moon from the extended distribution of thorium near the Compton-Belkovich Volcanic Complex

Abstract

We reconstruct the abundance of thorium near the Compton-Belkovich Volcanic Complex on the Moon, using data from the Lunar Prospector Gamma Ray Spectrometer. We enhance the resolution via a maximum entropy pixon image reconstruction technique, and find that the thorium is distributed over a larger (40 km $\times$ 75 km) area than the (25 km $\times$ 35 km) high albedo region normally associated with Compton-Belkovich. Our reconstructions show that inside this region, the thorium concentration is 15$-$33 ppm. We also find additional thorium, spread up to 300 km eastward of the complex at $\sim$2 ppm. The thorium must have been deposited during the formation of the volcanic complex, because subsequent lateral transport mechanisms, such as small impacts, are unable to move sufficient material. The morphology of the feature is consistent with pyroclastic dispersal and we conclude that the present distribution of thorium was likely created by the explosive eruption of silicic magma.

3.1 Introduction

The chemical composition of the Moon's surface was mapped by the Lunar Prospector spacecraft using gamma ray and neutron spectroscopy (Elphic et al., 1998, 2000; Feldman et al., 1998, 2000, 2002; Lawrence et al., 1998a, 2000, 2002; Prettyman et al., 2006) and these maps have led to an improved understanding of the formation and evolution of the

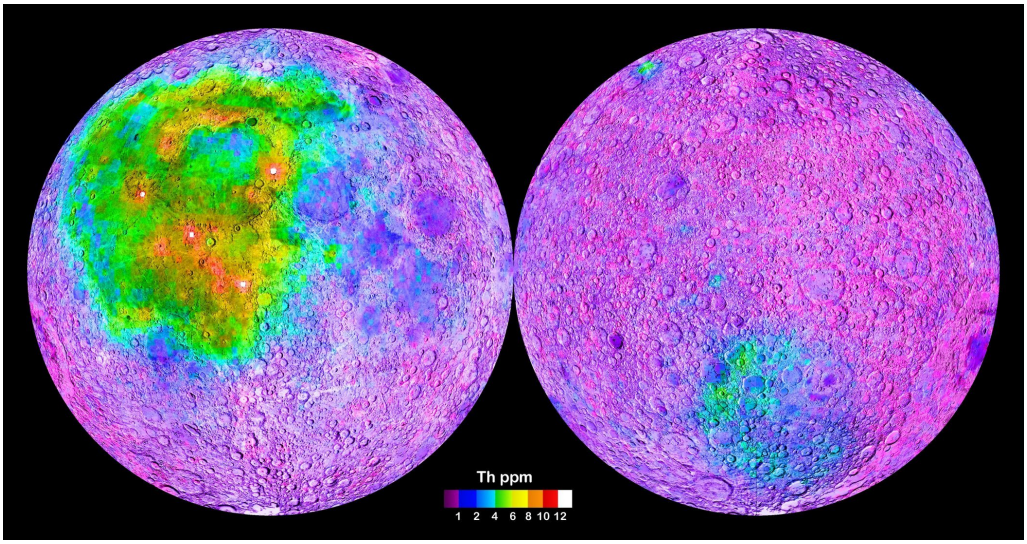

Figure 3.1: The lunar nearside (left) and farside (right) Th distribution from Lunar Prospector Gamma Ray Spectrometer data. The Compton-Belkovich Th anomaly is seen at the top-left of the right hand panel, with Th abundances approaching 6 ppm. Image credit: NASA/JPL.

lunar surface and interior (Jolliff et al., 2000; Hagerty et al., 2006, 2009). In the case of gamma ray detection from the natural decay of thorium (Th), uranium and potassium, the inferred abundances do not depend on cosmic ray flux or ground truth but only on having an accurate background subtraction (Metzger et al., 1977). The global distribution of Th is shown in figure 3.1. These chemical elements are particularly interesting as they are large and incompatible so preferentially partition into the melt phase during magmagenesis, and remain in the melt phase as it crystallizes. Thus the distribution of these elements acts as a tracer of magmatic activity and differentiation.

Of the three chemical elements detectable from orbit, Th is the most easily observed because its 2.61 MeV peak in the Moon's gamma ray spectrum is both strong and well separated from other peaks (Reedy, 1978). Examination of Th abundance maps, along with other data, gave rise to the interpretation that the lunar surface comprises three terranes (Jolliff et al., 2000): the low-Th Feldspathic Highland Terrane; the moderate-Th South Pole-Aitken basin; and the high-Th Procellarum KREEP Terrane (named after the materials with high potassium (K), rare earth element (REE) and phosphorus (P) abundances that cover much of its surface but which also contain other incompatible elements including Th (Warren & Wasson, 1979)).

Several anomalous regions of the Moon's surface fall outside this broad classification scheme, most notably a small but distinct Th enrichment located between the craters Compton (55.3°N, 103.8°E) and Belkovich (61.1°N, 90.2°E) on the lunar farside (which can be seen in the upper left of the right panel of figure 3.1). An isolated enrichment of Th was first detected at (60°N, 100°E) in the Lunar Prospector Gamma Ray Spectrometer (LP-GRS) data (Lawrence et al., 1999, 2000; Gillis et al., 2002). Jolliff et al. (2011b) associated this compositionally unique feature with a $\sim$25 km $\times$ 35 km topographically elevated, high albedo region that contains irregular depressions, cones and domes of varying size. They interpreted this region as a small, silicic volcanic complex, which they referred to as the Compton-Belkovich Volcanic Complex (CBVC) (Jolliff et al., 2011b). Uniquely, the high Th region of the CBVC is not coincident with an elevated FeO terrain (as would be expected from a KREEP basalt); instead the CBVC appears to have a low FeO abundance ($\sim 4-5$wt.%) that is similar to much of the lunar highlands (Lawrence et al., 1999; Jolliff et al., 2011b). Crater counting results indicate a likely age greater than 3 Ga for volcanic resurfacing at the CBVC (Shirley et al., 2013), suggesting that the Th distribution exposed at the CBVC may provide a rare insight into the extent of fractionation and the distribution of such magmatic activity at this time in the Moon's evolution.

One drawback of gamma ray spectroscopy is the large spatial footprint of a gamma ray detector. When the LP-GRS was in an orbit 30 km above the lunar surface, the full-width at half-maximum (FWHM) of the detector's footprint was $\sim$ 45 km (Lawrence et al., 2003). Additional statistical analysis is therefore required to extract information about the chemistry of sites as small as the CBVC. In this chapter we use the maximum entropy pixon method (Pina & Puetter, 1993) to remove blurring caused by the large detector footprint and enhance the spatial resolution of gamma ray data (by a factor of $1.5-2$ compared with other image reconstruction techniques; Lawrence et al. 2007), in a way that is robust to noise. This allows us to test the prevailing hypothesis regarding the distribution of Th around the CBVC — that it is all contained within the high albedo region (Jolliff et al., 2011b). Under that assumption, the raw counts from the LP-GRS data imply a Th concentration within the feature of $\sim 40-55$ ppm (Lawrence et al., 2003), which is important because only one known lunar rock type has such high concentrations of Th, namely granite/felsite (Jolliff, 1998).

3.1.1 Lunar volcanism

Basaltic volcanism was once common on the Moon and is responsible for the lunar mare that cover 17% of the lunar surface (Head, 1976), mostly filling the near-side basins. Evidence for basaltic, non-mare volcanism is most evident in the dark glasses that are distributed across the lunar surface, which are thought to be the product of basaltic pyroclastic eruptions.

Silicic, non-mare volcanism is much less common, observed at only a handful of locations including Hansteen alpha (Hawke et al., 2003), Mairan hills, Lassell Massif (Hagerty et al., 2006), the Gruithuisen Domes (Chevrel et al., 1999) and Compton-Belkovich (Jolliff et al., 2011b). These silicic volcanic constructs are steep-sided, with widths of a few km and heights greater than 1 km. Their morphology, enhanced Th concentrations and Christiansen features (emissivity maxima associated with SiO stretching vibrations (Salisbury et al., 1970)) all imply that these features are the result of evolved, silicic volcanism (Hagerty et al., 2006; Glotch et al., 2010). All of these features are located within the Procellarum KREEP Terrane, except the CBVC, which is on the lunar farside. One explanation of the origin of the silicic domes is that they are formed by the eruption of magma that is produced when ascending diapirs of basaltic magma stall at and underplate the base of the crust, causing it to partially remelt; the resulting melt is more silicic than the original basalt, and is enriched in incompatible elements and phases (Head et al., 2000). Alternatively the basaltic magma might stall at the base of the megaregolith, then undergo fractionation to produce a more evolved, silicic magma (Jolliff et al., 2011a), which subsequently erupts.

Silicic, non-mare volcanic centres have previously been assumed to be similar in nature to terrestrial rhyolite domes (Hagerty et al., 2006), which are erupted extrusively. Jolliff et al. (2011b) suggest that pyroclastic material may have been distributed over distances of a few kilometres from the CBVC but, to our knowledge, no evidence has previously been presented for vigorous explosive silicic volcanism on the moon.

3.1.2 Th rich minerals at the CBVC

The association of high Th concentrations in lunar samples with granite or felsite is clear, with Th concentrations of granitic samples generally falling in the range $20-65$ ppm (Seddio et al., 2013). Granitic assemblages clearly form from highly differentiated melt compositions that are enriched in many of the incompatible trace elements, especially the large-ion-lithophile (LIL) elements. Mafic evolved assemblages also occur in the lunar samples that exhibit LIL enrichment, such as alkali anorthosite and monzogabbro (Jolliff, 1998); however, these assemblages do not contain as high Th and U concentrations as do some of the granitic samples. Alkali anorthosites have Th concentrations as high as 40 ppm, but most have < 20 ppm, and monzogabbro samples have Th concentrations as high as about 45 ppm (Wieczorek et al., 2006), but they have substantially higher FeO ($\sim 10-16$ wt.%) than is indicated for the CBVC by LP-GRS data (Jolliff et al., 2011b). KREEP basalts only contain up to about 15 ppm Th and FeO typically in excess of 10 wt.%.

3.2 Methods

We will use the maximum entropy pixon method, described in section 2.5.1, to improve the resolution and suppress noise in the LP-GRS data. Following a brief discussion of the data we use, our assumptions about both the point spread function and noise will be detailed in section 3.2.2. The modifications to the maximum entropy pixon method that we will use to deal with the small size of the CBVC are described in section 3.2.2.

3.2.1 Data

Time-series LP-GRS observations from the 7 months that the Lunar Prospector spent at an altitude of ~ 30 km are used in this work. Each observation accumulated a gamma ray spectrum over an integration period of $\Delta t = 32$ s, giving 490952 observations in total. The reduction of these data was described by Lawrence et al. (2004), with the counts in the Th decay line at 2.61 MeV being defined as the excess over a background value within the $2.5-2.7$ MeV range. Absolute Th abundances are inferred following the procedures

used by Prettyman et al. (2002b) and Prettyman et al. (2002a). The typical count rate is $R = 3.7$ gamma rays per second, with the conversion to Th concentration in ppm being

$$[\mathrm{Th}] = 4.02R - 14.05, \tag{3.2.1}$$

from Lawrence et al. (2003).

3.2.2 Assumed instrumental properties

We begin with the LP-GRS PSF model by Lawrence et al. (2004), which is circular and has a FWHM of ~ 45 km, but we modify this to take into account the motion of the spacecraft. The detector moved at $\sim 1.6\,\mathrm{kms}^{-1}$ with respect to the lunar surface, i.e. ~ 51 km during the 32 s integration period. We convolve the circular PSF with a line extended in the direction of motion of the spacecraft, with length equal to the distance traveled by the spacecraft during one observation. This produces an elliptical PSF that is elongated in the direction of the poles.

As the LP-GRS was a counting experiment, the number of Th decay gamma rays received above a particular patch of lunar surface should follow a Poisson distribution. However, during the data reduction process, corrections have been applied such that the reduced count rates could have a somewhat different distribution. These corrections compensate for temporal variations in the galactic cosmic ray flux, the varying altitude and latitude of the spacecraft and the detector dead-time (i.e. the interval after a detection in which another cannot be registered) (Lawrence et al., 2004). Despite these various non-negligible corrections, Lawrence et al. (2004) showed that the noise on the Th-line data was surprisingly close to Poisson. In order that the reconstruction method can appropriately weight each observation, it is important to understand the statistical properties of the noise.

To gauge the effects of these corrections on the data we created a mock set of data in which the noise was known to be Poisson and to which each of the corrections (for temporal variations in the galactic cosmic ray flux, the varying altitude and latitude of the spacecraft and the detector dead-time, described in detail in Lawrence et al. (2004)) were applied in turn. This mock data set was created by:

1. undoing the corrections made to the Th-line LP-GRS time-series observations to find the measured number of counts in each observation;

2. making a time-series of Poisson random variables with means equal to the number of counts found in the previous step;

3. applying the four corrections to create the mock time-series.

Once the mock time-series was created, the statistics were tested by binning the data in $2^\circ \times 2^\circ$ pixels. This binning is the finest possible contining multiple repeat observations in all pixels. The expected scatter in observations within each of these pixels is well approximated by a Gaussian with width $\sigma = \sqrt{R\Delta t}$. This standard deviation was compared with that measured from the repeat observations of the same pixel. They were found to be in good agreement, with the average ratio of the two differing from unity by less than 1%.

It so happens that the correction of count rates to an altitude of 100 km typically decreases the corrected count rate by $\sim 15\%$ whereas the normalisation of the background to the high initial galactic cosmic ray flux increases the corrected count rate by $\sim 10\%$ and correcting for deadtime increases it by $\sim 4\%$. Consequently, these factors approximately cancel and the variance of the corrected measurements can be accurately treated as equal to the mean.

Decoupling

The pseudoimage smoothing in the maximum entropy pixon method described in the previous chapter makes this algorithm unable to produce sharp boundaries in a reconstruction. If such boundaries are demanded by the data then the residuals will be large, indicating that the reconstructed image represents a bad fit to the data. For the basic pixon reconstruction of the LP-GRS Th data in the region of the CBVC, such a problem occurs. The red line in the left panel of Fig. 3.2 shows the radial variation of the residuals from the point with highest Th, when using a basic pixon reconstruction. The positive residuals at small separations reflect an underestimation of the central reconstructed Th abundance resulting from the pixons oversmoothing this part of the reconstruction.

To incorporate very high spatial resolution features in the reconstruction, where we have prior information that they may exist (in this case provided by Jolliff et al. (2011b) in the form of Diviner data and optical imaging) we exploit a technique described by Eke et al. (2009). Within a region of the image marked out by high residuals and prior information, which we shall call the "decoupled region", we introduce a separate pseudoimage that does not affect the reconstruction outside the decoupled region (and vice versa). For the reconstructions in this chapter, the decoupled region is effectively the high Th region including the CBVC.

Using a decoupled rectangular area of size 50 km $\times$ 70 km centred on the CBVC, leads to the results as shown with blue curves in Fig. 3.2. In addition to removing the non-zero radially-averaged residuals, as shown in the left hand panel, the posterior probability in the right-hand panel also reflects how decoupling dramatically improves the reconstruction. This demonstrates that the new technique gives rise to an inferred truth, i.e. reconstruction, that is more consistent with the data.

Unfortunately, the posterior probability curves are too noisy for their maxima to be easily located in an automated way. Indeed, this noise is further increased when using decoupling, because the image optimization algorithm converges to fits in which the total Th abundance in the decoupled region varies. We therefore chose to fix the count rate within the decoupled region, then scan a range of possible count rates for each step in the posterior maximization. We can, in any case, say little about how the count rate varies within the decoupled region, because the instrumental PSF is comparable to the size of the decoupled region. Nonetheless, we have explicitly tested distributions of Th within the decoupled region that decrease with distance like a Gaussian, and these are slightly disfavoured by the data.

3.3 Results

We vary the size, shape, position and Th content of the decoupled region to determine the optimum reconstruction for the Th data in the vicinity of the CBVC. The resulting, unblurred image is shown in Fig. 3.3. A high Th region that is 50 km $\times$ 70 km, i.e. larger than the high albedo area, and a more extended lower concentration Th zone are both

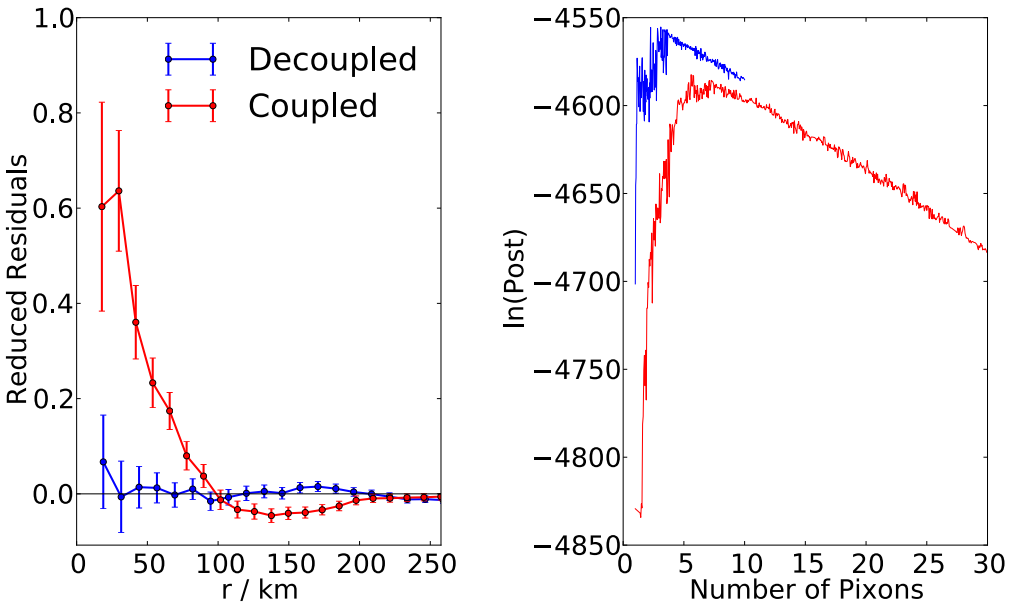

Figure 3.2: *Left panel:* mean residuals as a function of radius from the centre of the CBVC. *Right panel:* the posterior probability of the reconstruction as a function of the number of pixons used. In both panels, red (blue) points show the performance without (with) decoupling. The corresponding, decoupled, inferred 2D Th abundance is shown in Fig. 3.3.

required by the data. In this section, we will discuss the statistical significance of these features. Our discussion divides naturally into considerations of the high Th region in the vicinity of the CBVC and the more extended spatial distribution of Th. We consider the geological implications of our results in section 3.4.

3.3.1 The high Th region

Lawrence et al. (2003) considered the Th excess at the CBVC to be localized within the $\sim 25\,\mathrm{km} \times 35\,\mathrm{km}$ high albedo region, at an abundance of $\sim 50\,\mathrm{ppm}$. We have tested this hypothesis by examining a set of reconstructions that spread the Th excess across different sized and shaped high Th regions. While the rectangular high Th region shown in Fig. 3.3 is slightly favoured over a circular one, the difference is not large. Thus, to reduce the size of the parameter space to be tested, a circular high Th region centred on the middle of the high albedo region is adopted when considering the effects of changing the area of the high-Th region. This choice leaves only two free parameters; the area of the high-Th region and the Th concentration within it. The relative merits of these different choices

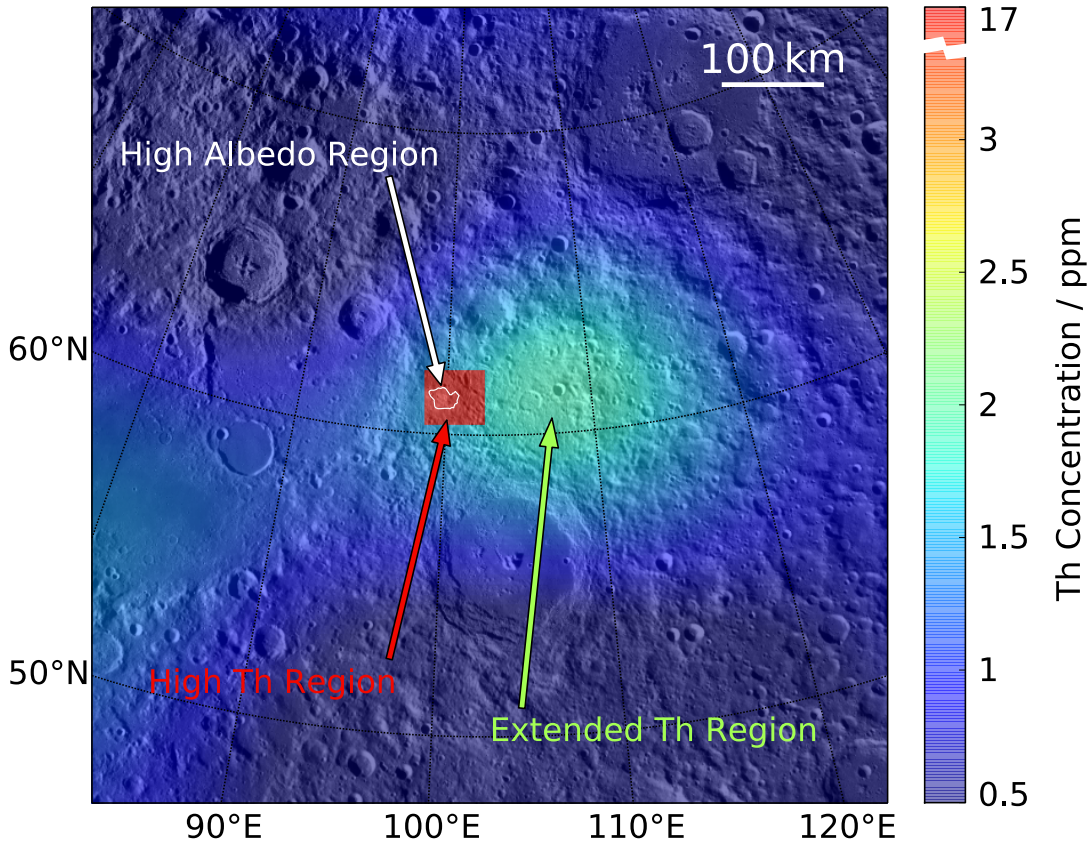

Figure 3.3: The best-fitting, unblurred Th distribution in the vicinity of the CBVC. The shape and size of the decoupled region shown was found to be more consistent with the data than others. The range of Th concentration is truncated to increase the contrast in the coupled, low Th region. The high Th region is shown in red and occupies approximately 5 times the area of the high albedo feature, which is shown with the white contour. Underlayed is a Lunar Reconnaissance Orbiter Wide Angle Camera image of the area around the CBVC.

are quantitatively assessed using the misfit statistic

$$\chi^2 = \sum_{i=1}^{m} R(i)^2, \tag{3.3.2}$$

defined in the m pixels within 60 km of the centre of the CBVC (this region focuses on the area where the highest Th count rates are concentrated, and is necessarily broader than the instrumental PSF). The misfit statistic is driven mainly by the size of the decoupled region, rather than its shape or precise location.

The position of the black star in Fig. 3.4 indeed indicates a preferred concentration of 51 ppm if the Th distribution is constrained to be within the 25 km $\times$ 35 km high albedo

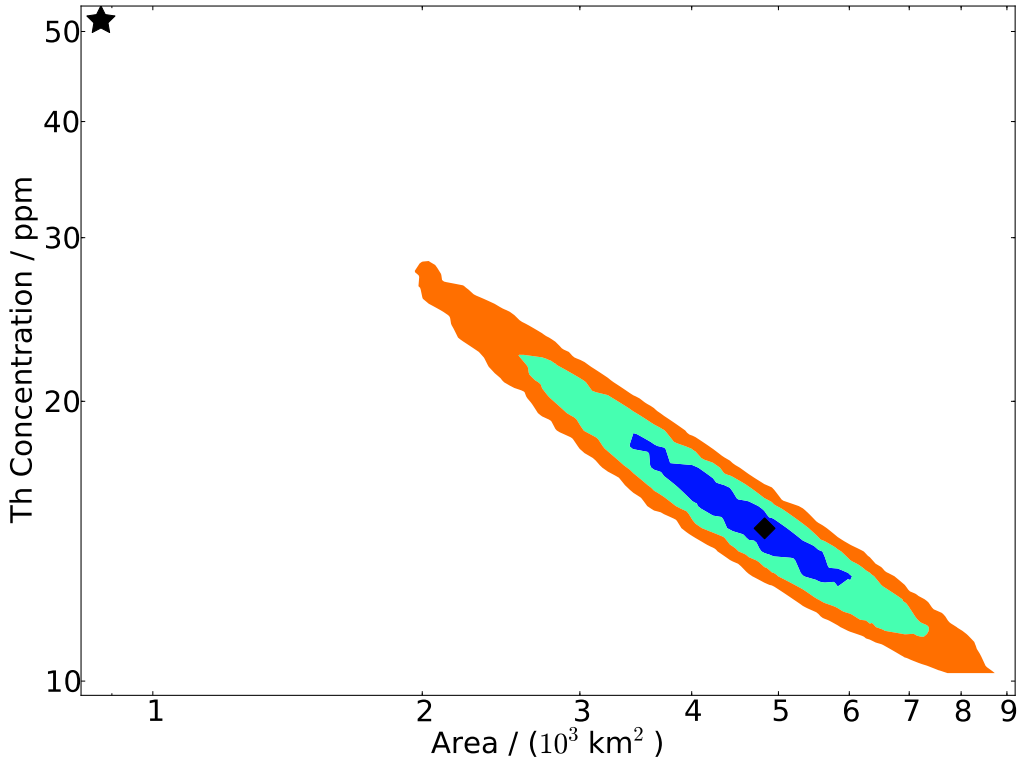

Figure 3.4: Constraints on the size of the high Th region in reconstructed images, and the Th concentration inside that region. The black diamond shows the best-fit reconstruction with minimum χ^2. Coloured regions enclose 68%/95%/99.7% confidence limits, determined using $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$. The black star shows the optimum reconstruction under the assumption that the decoupled region coincides with the high albedo feature.

region. However, such a concentrated distribution of Th is very strongly disfavoured. The LP-GRS data are much better fit by reconstructions in which the Th is uniformly distributed over 3000–5800 km^2, corresponding to side lengths of $\sim$ 55–75 km, at lower concentrations of $\sim$ 20–13 ppm. This area is approximately 5 times larger than the high albedo feature, but still only slightly bigger than the LP-GRS PSF, so our measurement of its area is necessarily imprecise.

3.3.2 The extended Th region

Outside the central high-Th region there are two regions with enhanced Th content, the first, to the WSW, has a Th content less than 2 ppm and is coincident with the eastern edge of Mare Humboldtianum and so is not directly related to the silicic CBVC. The second has a Th content up to 2.7 ppm and extends $\sim$ 300 km east from the CBVC. This

feature is evident in Fig. 3.3, in the results from the forward modelling in Lawrence et al. (2003), and in the raw data (Jolliff et al., 2011b). We will refer to it as the extended Th region. However, any noisy image reconstruction may contain spurious features that are not demanded by the data. We shall therefore assess the statistical significance of this extended enhanced Th region, by determining the probability that a similar excess would have been reconstructed by chance, even if it had not actually been present in the underlying map.

We simulated 1200 mock data sets of a model with no extended Th excess; the model Th concentration is high inside the decoupled region, and low outside (equal to the mean of pixels outside the decoupled region in Fig. 3.3). Each mock data set is created by blurring this model map with the instrumental PSF, then taking a different noisy realisation to give the observed gamma-ray count rates. These simulated data are then reconstructed, producing 1200 Th count rate values in each map pixel, from which we can work out the probability of any false positive reconstructed excess being found in the actual LP-GRS Th reconstruction. Fig. 3.5 shows the distributions of simulated count rates for the three pixels labelled in Fig. 3.6. Also shown are the best-fitting Gaussian curves through the mock count rate distributions. These curves describe the results well and allow us to extrapolate to probabilities outside the range accessible with only 1200 samples. Pixel 1 lies in a region where the mean reconstructed model background count rate matches that in the reconstruction of the LP-GRS data, whereas the LP-GRS data have a higher count rate in pixel 2 and lower one in pixel 3.

In each pixel, one can calculate the probability that the reconstructed mock Th count rate is lower than that obtained from the reconstruction of the LP-GRS data. This allows us to test the null hypothesis underpinning the mocks, namely that the excess Th is concentrated entirely into the high Th region. Doing so yields the results shown in Fig. 3.6. Regions of Fig. 3.6 with probabilities close to either 0 or 1 should be considered unlikely. The existence of such regions suggests that there are significant variations in the Th concentration outside the CBVC. One statistically significant extended zone of excess Th reaches up to $\sim 300\,\mathrm{km}$ east of the CBVC.

As a simple check of our procedure, we also compute the statistical significance of the excess counts in the $50\,\mathrm{km} \times 50\,\mathrm{km}$ square region centered 120 km east of the CBVC,

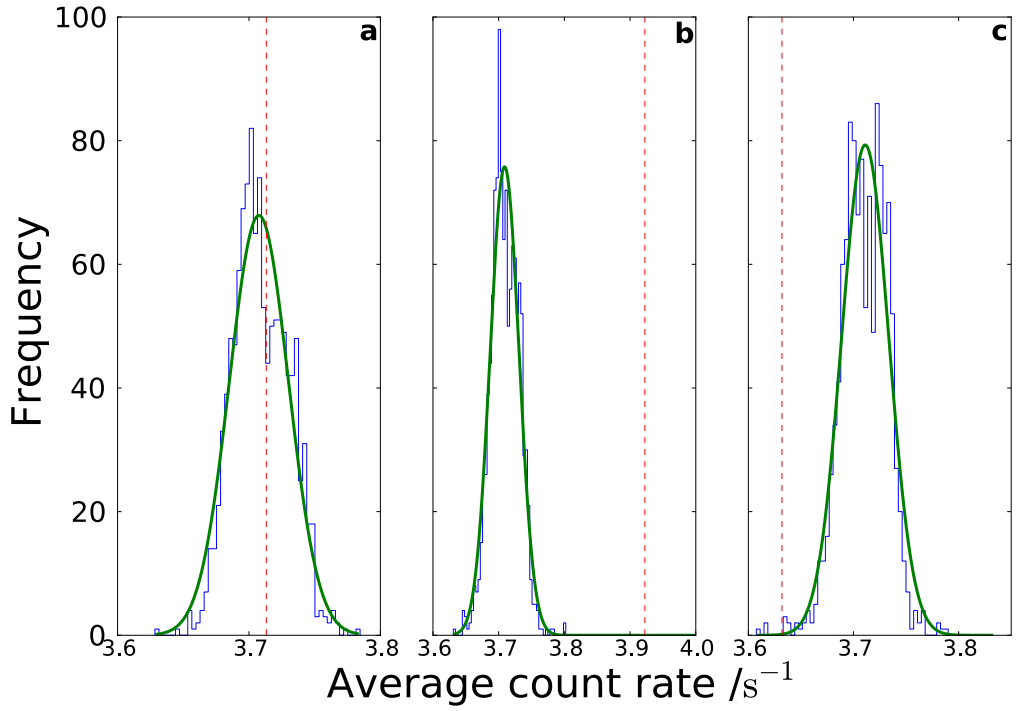

Figure 3.5: Comparison of the reconstructed Th count rate from real LP-GRS data, and from mock data assuming no extended Th features, in 3 different map pixels (marked in Fig. 3.6). The distributions from the mock data sets are the blue histograms, with best-fitting Gaussians shown with green curves. In each panel the vertical dotted red line is the count rate in the corresponding pixels in the reconstruction of the LP-GRS data. The probability of obtaining less than the predicted count rate can be found by integrating the fit to the count rate distribution up to the black line.

which is sufficiently distant from the high Th area to receive few counts as a result of the PSF blurring. This area has a 5σ excess in counts, strongly suggesting that the extended Th region to the east of the CBVC is a statistically significant Th excess.

3.4 Implications for the origin of the Th distribution

The results in the previous section imply that the high Th region is larger than the $\sim$ 25 km $\times$ 35 km area of silicic composition identified in Diviner data and the area of increased reflectance identified in the Wide and Narrow Angle Camera (WAC, NAC) imaging (Jolliff et al., 2011b). This result might imply that the Th was emplaced in the high albedo region and has subsequently undergone lateral transport to produce the current distribution, or that the process that placed the Th on the lunar surface itself imprinted this

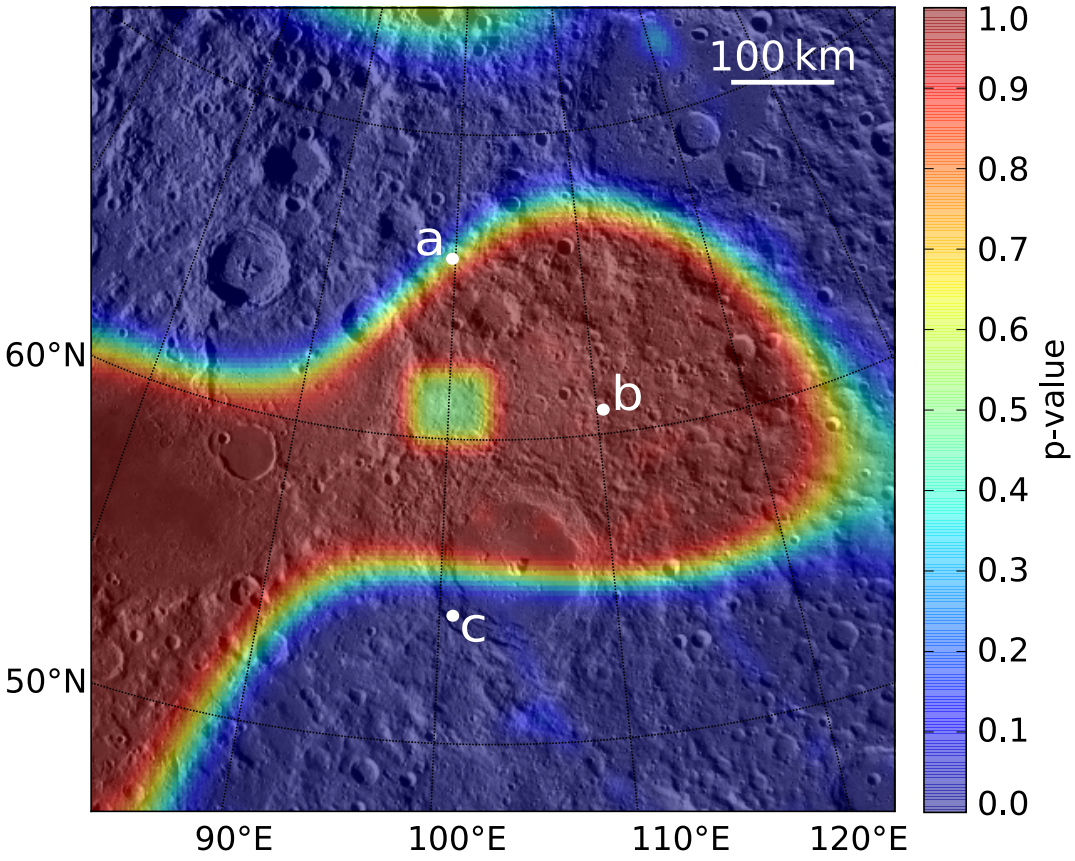

Figure 3.6: The probability that the predicted count rate is less than that obtained from the reconstruction of the LP-GRS Th-line data set assuming the null hypothesis to be true in each pixel. The results have been smoothed with a Gaussian kernel with FWHM equal to that of the PSF in order to suppress noise. Underlayed is a Lunar Reconnaissance Orbiter Wide Angle Camera image of the area around the CBVC.

extended distribution.

Assuming that the high Th material was initially emplaced within the CBVC via silicic volcanism, as proposed by Jolliff et al. (2011a), then the high albedo region can be taken to trace the original extent of the Th on the surface – leaving its subsequent transport to be explained. Another possibility is that the original Th distribution was not coincident with the high albedo region and that the presence of the regions with elevated Th contents outside the albedo feature was caused by pyroclastic eruptions at a time close to or at the formation of the volcanic complex. This hypothesis was proposed by Jolliff et al. (2011b) to explain the eastward extension of the Th distribution beyond the high albedo region over distances of $\sim 7\,\mathrm{km}$.

3.4.1 The sputtering of Th atoms

Sputtering liberates atoms from the surface of the lunar regolith, but most sputtered atoms have speeds greater than the escape speed (Wurz et al., 2007) so sputtering tends to remove material altogether. However, the most probable speed of a sputtered particle with mass m is expected to scale with $m^{-\frac{1}{2}}$ (Wurz et al., 2007). As Th atoms have $m = 232$ amu the typical speed of a sputtered Th atom is $\sim 0.8\,\mathrm{kms}^{-1}$, considerably less than the escape speed from the Moon ($\sim 2.4\,\mathrm{kms}^{-1}$). Using the model by Cassidy & Johnson (2005) for the distribution of polar angles, α, of the sputtered atoms,

$$f(\alpha) \propto \cos(\alpha), \tag{3.4.3}$$

and assuming that the azimuthal angular distribution is uniform, we can find the lateral velocity of sputtered atoms and hence the average distance travelled by the atoms before they fall back to the lunar surface, d. This is done by averaging over polar angle the product of the time of flight and lateral velocity. Assuming for simplicity that the orbit is parabolic, which turns out to be sufficiently accurate,

$$\langle d \rangle = \int_0^{\frac{\pi}{2}} \frac{2v^2}{g} \cos^2(\alpha) \sin(\alpha) d\alpha, \tag{3.4.4}$$

$$= \frac{2v^2}{3g}, \tag{3.4.5}$$

where v denotes the most likely initial speed of the sputtered atoms and g is the acceleration due to gravity. Using the most likely value of $v = 0.8\,\mathrm{kms}^{-1}$ gives $d = 250\,\mathrm{km}$.

The equations above characterise the average hop of a sputtered Th atom. However, to find the impact that this process has on the concentration of Th in the vicinity of the CBVC we also need the rate of sputtering. A rough estimate of this can be made by ignoring binding energy variations and assuming that the number of atoms of a particular species that are sputtered from the regolith is proportional to the number density of atoms of that species in the regolith. Taking the sputtered flux of oxygen given in Wurz et al. (2007) (where an average solar wind ion flux of $4.5 \times 10^{12}\,\mathrm{m}^{-2}\mathrm{s}^{-1}$ is assumed) and an oxygen concentration of 46wt.% (Heiken et al., 1991) in the lunar regolith, versus a typical Th concentration from our reconstructions of 20 ppm, we estimate the flux of sputtered Th

atoms to be $\Phi_{\mathrm{Th}} \sim 4 \times 10^5\,\mathrm{m}^{-2}\mathrm{s}^{-1}$. The average time a Th atom would spend on the surface before being sputtered is then given by

$$\tau_{\mathrm{sputter}} = \frac{n_{\mathrm{Th}} d_{\mathrm{sputter}}}{\Phi_{\mathrm{Th}}}, \tag{3.4.6}$$

where n_{Th} is the volume number density of Th atoms, which is related simply to the Th concentration and regolith density, and $d_{\mathrm{sputter}} = 10\,\mathrm{nm}$ is the depth of regolith susceptible to sputtering (Barghouty et al., 2011). Using the values from above gives $\tau_{\mathrm{sputter}} \sim 130\,\mathrm{yr}$.

The final step in the consideration of this process is to assess the effect of the overturn of regolith on the concentration at the surface. We assume that the rate of overturn of regolith due to gardening is constant and $\sim 0.5\,\mathrm{m/Ga}$ (Hörz, 1977). If this gardening occurs at a constant rate, then the time that Th atoms spend on the surface available for sputtering is $\tau_{\mathrm{overturn}} \sim \tau_{\mathrm{sputter}}/100$. This implies that a conservative approximation, when trying to determine the maximum amount of Th that could leave the CBVC this way, is to consider the Th at the surface of the CBVC to be constantly renewed by the overturn of regolith. Therefore the effect of sputtering on the CBVC is that, every year, approximately one in ten atoms in the upper 10 nm are sputtered and move away from the CBVC with a typical step size of $\sim 250\,\mathrm{km}$. The sputtered atoms that are re-implanted in the regolith are then most likely gardened down and never take another hop. The net effect over the $\sim 3.5\,\mathrm{Ga}$ of the CBVC's lifetime is that $\sim 10\%$ of the Th atoms in the top $\sim 1\,\mathrm{m}$ of regolith, the region accessible to the LP-GRS data, will have left the CBVC and settled in the surrounding few hundred km. This dispersal would increase the Th concentration in the area surrounding the CBVC by considerably less than 1%, which is not enough to explain the findings of section 4.3.1.

It should be noted that the above argument places an upper limit on the effect of sputtering on the concentration at the CBVC because of the assumptions that the overturn of regolith occurs continuously and that material, once gardened from the surface, is randomly distributed throughout the underlying regolith. In practice neither of these assumptions hold exactly. Gault et al. (1974) suggested that the main cause of gardening is impacts of small meteorites and consequently it is only the upper mm of regolith that is continuously reworked and regolith deeper than 1 cm is rarely brought to the surface

(Hörz, 1977). Additionally if overturn is due, primarily, to micrometeorites it is best to think of overturn taking place to a depth of approximately a μm every kyr and not as a continuous process. We would, in this case, expect of order 0.1% Th to be lost from the CBVC instead of the $\sim$ 10% found above.

3.4.2 Mechanical transport of Th-bearing regolith

Meteorites impacting on the lunar surface cause lateral mixing of regolith. When the regolith is made up from two compositionally distinct components this lateral transport can lead to a diffusion-like effect in which the two regolith types are mixed mechanically. The bulk regolith composition at any point is a weighted average of the two end states. We have calculated the effect of this process on the Th concentration at the CBVC using the model described in Li & Mustard (2005) and Marcus (1970) under the assumption that the CBVC was originally a compositionally homogeneous, circular feature 35 km in diameter surrounded by a uniform background.

The lateral transport model assumes the following power law relationships for the number of craters, T, above a given crater diameter, x_c, and the ejecta thickness, ζ, with distance from the crater rim, r,

$$T = F x_c^{\gamma}, \tag{3.4.7}$$

$$\zeta(r) = R_0 x_c^h \left(\frac{x_c}{2r}\right)^k \tag{3.4.8}$$

where F, R_0, h and k are constants. We set the constants of the crater rim ejecta height (R_0 and h) using the data from Arvidson et al. (1975) as these are thought to be relevant for small impacts ($x_c < 1$ km) and it is presumably the frequent, smaller impacts that contribute most to the dispersal of high Th material from the CBVC. We take $F = 2.5 \times 10^{-3}\,\mathrm{km}^{-2}$ as is appropriate for late Imbrian ages (Wilhelms et al., 1987). The only constraints we place on k are those theoretical limits suggested by Housen et al. (1983), that $2.5 < k < 3.0$. Consequently the model requires γ to lie in a certain range as it must obey the condition

$$2/k < \gamma/(k+h). \tag{3.4.9}$$

Using this model and assuming cratering to be a random process allows one to derive a relationship between the total thickness of regolith, Z, at a particular location, that originated at least some given distance, ρ, away from that location. The formalism does not, however, give a value of the ejecta thickness at a particular point - only the characteristic function of the probability density function (p.d.f.) of the total ejecta thickness, $\phi_\rho(u)$:

$$\phi_\rho(u) = \exp -\lambda|u|^{\alpha_\rho}[1 - i\,\mathrm{sgn}(u)\tan(\pi\alpha/2)]), \tag{3.4.10}$$

$$\lambda = \frac{\pi\gamma R_0^{\alpha_\rho} F}{2\alpha_\rho k(\gamma - 2 - 2h/k)(2\rho)^{\alpha_\rho k - 2}} \times \Gamma(1-\alpha_\rho)\cos(\pi\alpha_\rho/2), \tag{3.4.11}$$

$$\alpha_\rho = \gamma/(k+h), \tag{3.4.12}$$

where Γ is the Gamma function, $i = \sqrt{-1}$ and $\mathrm{sgn}(u)$ is 1 if $u > 0$ and -1 if $u < 0$. Obtaining the p.d.f. from this characteristic function requires the use of a numerical integrator, and we follow Li & Mustard (2005) in using the STABLE code[1] (Nolan, 1999).

After the p.d.f. is obtained, the mode of the distribution is taken as the value of $Z(\rho)$. We combine these results with our model of the CBVC (that it was initially a circular feature 35 km in diameter) to find, as a function of distance from the CBVC, what fraction of the current regolith originated within the CBVC. The assumption is made that the regolith is well mixed and that the Th detected by the LP-GRS can be related directly to the proportion of ejecta at a particular point that originated from within the CBVC.

Fig. 3.7 shows the initial assumed Th concentration profile with a black solid line. The grey shaded region traces the variation, with distance, of the fraction of regolith that originated in the CBVC ~ 3.5 Ga earlier. A dashed line traces through the reconstructed Th map in the easterly direction, with the vertical scale chosen so that the value tends to zero at large distances and matches the model regolith fraction within the CBVC. While repeated small impacts are capable of moving some Th rich regolith away from the CBVC, it does not increase the Th concentration in the 5-15 km around the CBVC to the levels necessary to explain the difference between the high Th and high albedo regions implied by the results of section 4.3.1. The lateral transport model predicts that $\sim 25\%$

[1]The program STABLE is available from J. P. Nolan's website: academic2.american.edu/∼jpnolan

of the regolith currently within the CBVC originated from outside this region and has subsequently been redistributed into the CBVC by impacts. The effect of such transport would require that the measured Th concentration in the high Th region should be correspondingly increased in order to infer the initial concentration placed onto the surface by volcanism. Also shown in Fig. 3.7 is a curve showing the radial Th count rate variation in the reconstruction after it has been smoothed by the LP-GRS PSF, showing that the pixon method has increased the Th concentration in the high Th region by a factor of ~ 3 over that present in the blurred data.

In addition to the movement of regolith by impacts, one may hypothesise that, as the CBVC is a topographically elevated feature, down slope motion of regolith due to seismic shaking may be important in the lateral transport of Th bearing regolith. This is, however, not the case. The steepest slopes in the outer regions of the CBVC reach only $\sim 10^{\circ}$ from the horizontal, which, using the slope-distance relation from Houston et al. (1973) suggests lateral transport of ~ 1 cm in 3.5 Ga due to seismically induced downslope motion.

Effect of post-emplacement dispersal on Th content within the CBVC

Although the processes described in sections 3.4.1–3.4.2 did not greatly affect the Th concentration in the region surrounding the CBVC, lateral transport of regolith does have a significant effect on the measured abundance of Th within the CBVC. The fraction of regolith within the CBVC that originated there is between ~ 0.65 and 0.85 (Fig. 3.7). This suggests that the Th concentration when the CBVC was formed may have been $\sim 25-50\%$ greater than is detected today (approximating the surrounding regolith as having essentially a zero Th concentration). Consequently the Th concentration in the high Th region inferred from the reconstruction in section 3.3.1 underestimates that present when the material was emplaced onto the surface. As a result, the minerals that made up the CBVC at emplacement would have contained $\sim$ 15-33 ppm Th depending on the actual size of the feature and the parameters chosen in equation 3.4.7. We have not included the effect of sputtering on the change in concentration as our estimate is an upper bound and we suspect that the true contribution of this process is somewhat lower than 10%; however, if we were to include it, then this would raise the upper limit on the allowed Th

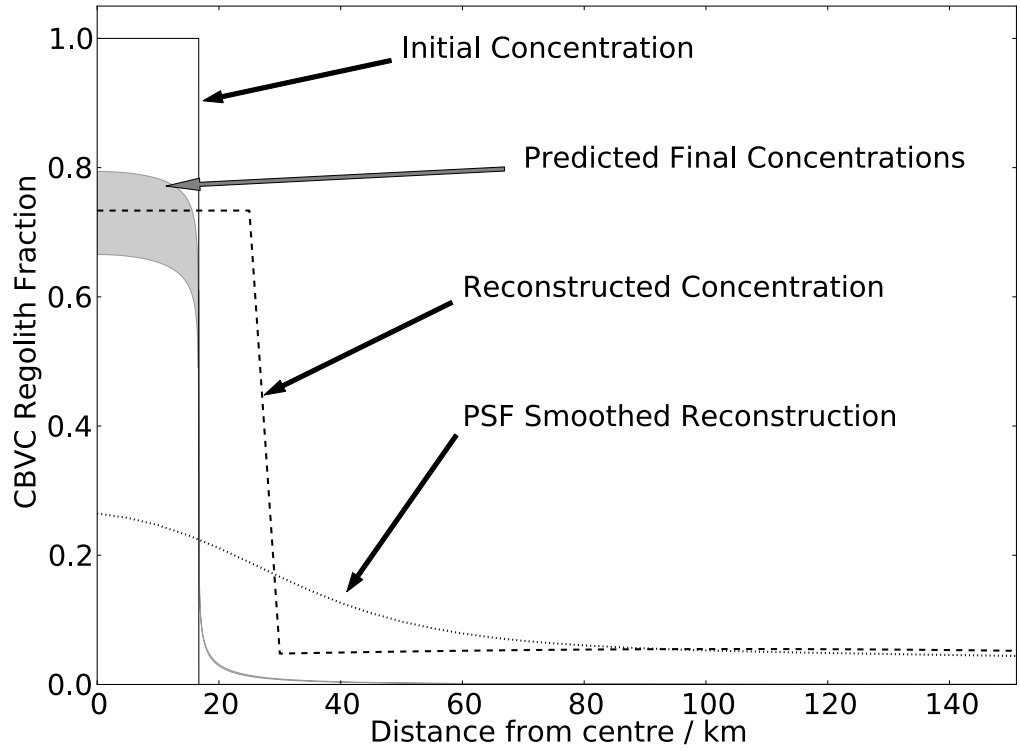

Figure 3.7: Assuming no pyroclastic dispersal, the variation of the fraction of regolith that originated within the CBVC with distance from the centre of the CBVC. The solid black line shows the hypothetical initial CBVC Th distribution, at time zero. The grey shaded area shows the range of solutions after 3.5 Ga of modification by impact processes. The dashed line is the present day reconstructed Th concentration (scaled so that it tends to zero at large distances and matches the model within the CBVC). The dotted line illustrates how the LP-GRS PSF suppresses the peak Th concentration being measured (scaled in line with the unsmoothed reconstruction).

concentration to 37 ppm.

The above Th concentrations along with the low FeO content around the CBVC imply that the rock components that are most likely to be present at the CBVC are granite/felsite and alkali anorthosite or some combination. In either case, the presence of alkali feldspar and a silica mineral (or a felsic glass) provides the best match for the LP-GRS Th and Fe data.

3.4.3 Lunar pyroclastic activity as a method of material transport

As has been shown in sections 3.4.1 - 3.4.2 the effects of post-emplacement processes to alter the distribution around the CBVC are insufficient to explain the extent of the Th distribution measured in the reconstructions in section 3.3.1. Therefore the Th must have been initially emplaced more widely than the high albedo region.

We hypothesise, following Jolliff et al. (2011b), that the mechanism was a highly silicic form of pyroclastic eruption not readily evident elsewhere on the Moon. Our results require dispersal over much greater distances of $\lesssim 300\,\mathrm{km}$, than proposed by Jolliff et al. (2011b). Repeated pyroclastic eruptions from the many volcanic features in the CBVC could feasibly give rise to the observed high Th regions that extend beyond the high albedo feature. The upper limit for ejecta distance for primitive pyroclastic eruption on the Moon is 350 km according to Wilson & Head (2003). One would expect that, as a melt evolves and becomes more concentrated in the volatile species that drives eruption, the ejection velocity would increase, implying that the range observed in the reconstructions is reasonable.

The total volume of material ejected from the CBVC that has given rise to the broad extended Th region to the east can be calculated, assuming that the ejecta had the same Th concentration as the CBVC in the reconstruction. The ejecta depth is

$$t = \frac{[\mathrm{Th}] - [\mathrm{Th}]_{\mathrm{background}}}{[\mathrm{Th}]_{\mathrm{ejecta}} - [\mathrm{Th}]_{\mathrm{background}}}, \tag{3.4.13}$$

where $[\mathrm{Th}]$ is the Th concentration at a given point, $[\mathrm{Th}]_{\mathrm{ejecta}}$ is the assumed Th concentration of the pyroclastic deposits and $[\mathrm{Th}]_{\mathrm{background}}$ is the Th concentration in the surrounding regolith. Integrating the ejecta depth over the feature gives an estimate of the total

ejecta volume of $8\,km^3$.

One may expect that the silicic material laid down during these pyroclastic events would be detected by Diviner; however, no spatial extension of the polymerized Christiansen feature position is seen much beyond the extent of the high albedo region (Jolliff et al., 2011b). This is readily explained since Diviner (and visible imaging) is sensitive only to the very surface composition whereas the LP-GRS is sensitive to a metre or so of depth. During 3.5 Ga of regolith gardening the silica emplaced by pyroclastic deposition, could have been mixed into the upper meter of regolith and effectively obscured from Diviner, whereas the Th signal would remain visible to the deeper-sensing GRS. This same argument applies to the non-detection of volatile rich material outside the CBVC in M^3 data (Bhattacharya et al., 2013; Petro et al., 2013).

3.5 Viability of lunar silicic pyroclastic volcanism

Typical lunar pyroclastic eruptions are driven by primitive magmas and give rise to dark coloured deposits (Head et al., 2002; Wilson & Head, 2003). The pyroclastic deposits we propose as the cause of the extension of the high Th region would be expected to have a high albedo because of their silicic (low-Fe) composition — akin to rhyolitic ash — which gives rise to a light-coloured assemblage of silica, alkali feldspar minerals and/or felsic glass. Smaller clast size also promotes higher albedo due to enhanced light scattering.

On Earth, explosive silicic volcanism produces abundant ash ($< 2\,mm$) and fine ash ($< 64\,\mu m$), which basaltic volcanism rarely does. Close to the vent, material is ejected as a jet, but the high surface area to volume ratio of the ash promotes rapid heat exchange with entrained atmosphere, producing buoyant, lofting plumes that may ascend tens of kilometres before attaining neutral buoyancy and spreading laterally (Sparks & Wilson, 1982). Winds then dominate the subsequent dispersal of the ash, which may circle the globe in the case of the largest eruptions. By contrast, on planetary bodies with negligible atmosphere such as the Moon, buoyant lofting is impossible. Instead, ejected particles will follow essentially ballistic trajectories, modified to some extent by particle–particle collisions.

We can calculate the maximum distance ejecta might be expected to travel in a lu-

nar pyroclastic eruption using the 1-dimensional gas flow equations. We treat the flow as a single-phase perfect gas and assume that ash particles will be accelerated to similar speeds to the gas. This is the standard approach in planetary science and has been verified experimentally (Kieffer & Sturtevant, 1984). Whilst more sophisticated multi-phase modelling is possible, our simple approach is sufficient to obtain rough estimates for attainable velocities.

We consider a gas of density ρ, temperature T, and velocity u moving steadily and essentially one-dimensionally along a conduit of cross sectional area A. For an ideal gas with adiabatic index γ, gas constant R_g, and specific heat at constant pressure c_p, we have the following relations between inlet (subscript 0) and outlet (subscript 1) conditions:

$$\frac{u_1}{u_0} = \frac{A_0}{A_1}\left(\frac{T_0}{T_1}\right)^{\frac{1}{\gamma-1}} \tag{3.5.14}$$

$$\frac{\rho_1}{\rho_0} = \left(\frac{T_1}{T_0}\right)^{\frac{1}{\gamma-1}} \tag{3.5.15}$$

$$\frac{p_1}{p_0} = \left(\frac{T_1}{T_0}\right)^{\frac{\gamma}{\gamma-1}}, \tag{3.5.16}$$

where $2c_p(T_0 - T_1) = u_1^2 - u_0^2$. Eliminating the cross sectional area, these equations can be rearranged to give the outflow velocity

$$u_1 = \sqrt{u_0^2 + 2c_pT_0\left[1 - \left(\frac{p_1}{p_0}\right)^{\frac{\gamma-1}{\gamma}}\right]}. \tag{3.5.17}$$

Expansion is limited by the condition that p_1 must be greater than or equal to the atmospheric pressure. Hence, if atmospheric pressure is significant, then this limits the exhaust velocity and leads to a complicated shock structure (Kieffer & Sturtevant, 1984). On the Moon however, p_1/p_0, is negligible and, assuming $u_1 \gg u_0$, we have a rough estimate for the largest attainable velocity

$$u_1 = \sqrt{2c_pT_0}. \tag{3.5.18}$$

The correct form for an imperfect gas is

$$u_1 = \sqrt{2\int_0^{T_0} c_p\,dT}. \tag{3.5.19}$$

For a temperature of 1100 K (a typical eruption temperature for a silicic magma on Earth) this gives a velocity of 1 430 m s^{-1}, where it has been taken that the gas will be predominantly carbon monoxide, which is produced in oxidation-reduction reactions between native carbon and metal oxides on nearing the surface (Fogel & Rutherford, 1995). This can be shown to be substantially greater than the launch speed necessary to emplace debris ballistically over 300 km, as is required by the reconstructions at the CBVC. The maximum range, measured along a great circle, for an object launched at speed v on an airless body of mass M and radius R is

$$x = 2R\sin^{-1}\left(\frac{Rv^2}{2GM - Rv^2}\right). \tag{3.5.20}$$

For the moon, where $g = 1.62\,\mathrm{m\,s^{-2}}$ and $R = 1.74 \times 10^6$ m, a speed of at least 669 m s^{-1} is necessary for a projectile to travel $x = 300$ km.

From equation (3.5.19), we calculate that a starting gas temperature of only 390 K is enough to achieve the required velocity of 669 m s^{-1} under perfect conditions. Of course many factors, including non-optimal vent orientation, will reduce these ideal ejection velocities but this simple calculation shows that it is straightforward for volcanic plumes on the Moon to eject material many hundreds of kilometres.

On Earth, volcanic conduits of all types may be inclined, and the vents are commonly asymmetric (Wood, 1980; Folch & Felpeto, 2005; Castro et al., 2013). For basaltic pyroclastic eruptions, which typically emplace ballistically, this may cause asymmetry in the distribution of pyroclastic material. For silicic eruptions, the rapid formation of a buoyant plume tends to disguise and overprint the effects of conduit inclination and vent asymmetry: the plume takes the ash straight up and wind is then dominant. In the absence of an atmosphere, silicic eruptions, too, would emplace ballistically, thus an inclined conduit would give rise to an asymmetric deposit. Furthermore, Jolliff et al. (2011b) note the presence of arcuate features in the topography of the CBVC that are suggestive of the collapse of volcanic edifices. Such collapses have been known to trigger directed, lateral blasts in silicic volcanoes on Earth, the 1980 eruption of Mt. St. Helens being the most famous example (Kieffer, 1981). It therefore seems plausible that the asymmetric distribution of pyroclastic material at the CBVC might have been caused by eruption from

an eastward-inclined conduit, or from an asymmetric vent open to the east, perhaps as a result of the collapse of a volcanic edifice.

Evidence for both effusive (dome-forming) and explosive (pyroclastic) eruption is seen at the CBVC. On Earth it is common for silicic volcanoes sometimes to erupt effusively, and sometimes explosively. This may be a consequence of variable differentiation of the melt, or variable composition of the magma at the point of formation (Sides et al., 2014), both of which may influence the viscosity and volatile content of the magma. The 2011 eruption of Cordón Caulle in Chile further demonstrated that silicic magma may simultaneously erupt effusively and explosively from a single vent (Castro et al., 2013). The variation in eruption style was inferred to result from variation in the path travelled by the magma during its ascent of the conduit, affecting its capacity to degas (Castro et al., 2013). An inclined conduit and asymmetric ash jetting — quickly masked by the formation of a buoyant plume — were also inferred for this eruption. All of these features suggest that, were it on Earth, the CBVC would not be so unusual a volcanic feature.

3.6 Conclusions

We have used the pixon image reconstruction method to produce the highest resolution map of the Th distribution around the Compton-Belkovich Volcanic Complex to date. This method largely removes the effect of the detector footprint from the Th map, in a way that is robust to noise present in the data. A central excess of Th had been previously assumed to be coincident with the 25 km × 35 km high albedo region observed in LP-NAC/WAC imaging (Lawrence et al., 2007; Jolliff et al., 2011b). However, we have shown that the central Th excess likely extends ∼40–75 km laterally, over a much larger area. The Th concentration in this region would have been ∼15–33 ppm at emplacement, with uncertainty driven by the precise current area and the amount of external Th-poor regolith that has been mixed into the CBVC during the past 3.5 Ga.

We identify an additional Th feature (significant at $> 5\sigma$), which extends ∼300 km east of the CBVC at a Th concentration of ∼2 ppm. The data outside the CBVC are certainly not consistent with a uniform low-Th background. The extended nature of the CBVC is not due to processes that have acted since its origin (e.g. lateral transport of

regolith and sputtering) so must have been present when it was formed. This distribution of Th is consistent with a mixture of pyroclastic eruptions, to distribute the Th widely, and effusive eruptions to produce the observed volcanic domes and high albedo region.

CHAPTER 4

Reconstruction of Mars Odyssey Neutron Spectrometer epithermal neutron data

In this chapter we present a map of the near subsurface hydrogen distribution on Mars, based on the Mars Odyssey Neutron Spectrometer epithermal neutron data but with much improved resolution. Via the first global Bayesian reconstruction of a remotely sensed planetary data set this map achieves a near two-fold improvement in linear spatial resolution compared with the instruments that make up the Mars Odyssey Gamma Ray Spectrometer suite, including the Neutron Spectrometer, Gamma Subsystem and High Energy Neutron Detector. The various image reconstruction techniques described in chapter 2 are applied to the mock data sets, similar to the MONS data, and the locally adaptive pixon method is found to be the most suitable for reconstructing MONS-like data.

4.1 Introduction

Knowing the present distribution of water in the Martian near-subsurface is important for several reasons: it allows inferences about the past and present climate to be drawn, which, in turn, give information about the dynamic history of Mars and the possibility the past, or present, existence of life. Additionally, understanding the small-scale distribution of water is important for landing site selection for missions looking for past or present signs of life or exploring in-situ resource utilisation (Squyres, 2011). The epithermal neutron flux from the Martian surface contains information about its composition, particularly the abundance of hydrogen, and therefore water, in the top few tens of centimetres. This information is extracted from the neutron data with the aid of Monte Carlo particle transport simulations, which require foreknowledge of the composition of the spacecraft

and the Martian surface and atmosphere. The application of image reconstruction techniques will improve the utility of the MONS data by removing noise, so that only features present on the surface will be present in the data, and improving resolution, thus allowing more features to be resolved.

The chapters following this will use the results of our reconstruction of MONS data to address a series of questions about the distribution of water and CO_2 on the surface of Mars. Before these questions can be answered we must understand the statistical properties of the MONS data and the instrument's PSF. As stated in chapter 2 these properties must be well understood if an image reconstruction technique is to be effective. Also, as we do not know that one image reconstruction technique is preferable for all data sets, the most effective image reconstruction technique for the MONS-like data must be determined.

4.2 Data

Approximately 3.5 Martian years of the time-series MONS prism-1 (Maurice et al., 2011) observations from JD 2452324.125388 (just before the start of Mars Year 26) until JD 2454922.577436 are used in this thesis. Prism-1 is the nadir facing detector on MONS and is most sensitive to epithermal neutrons. In this chapter and the next we restrict our attention to the 'frost-free' data, for which we use the same definition as Maurice et al. (2011) i.e. those data taken at a location (in time and space) in which the surface CO_2 frost area density in the NASA Ames GCM is less than 0.2 gcm^{-2}, in chapter 6 all of this data will be used.

The problem of choosing a pixelization scheme on the sphere is non-trivial as there is no analog of the natural, rectangular pixelization on the plane, in which all pixels have the same shape and size. For practicality, we require rows of iso-latitude pixels, with equal-area pixels within each row, as this reduces the time required for spherical harmonic deconvolution and synthesis (and therefore our image reconstruction calculations - see section 2.1.2). There are two common pixelization schemes that satisfy both of these criteria: the equirectangular pixelization, with pixels that are rectangular when plotted in an equidistant cylindrical (or equirectangular) projection, and HEALPix (Górski

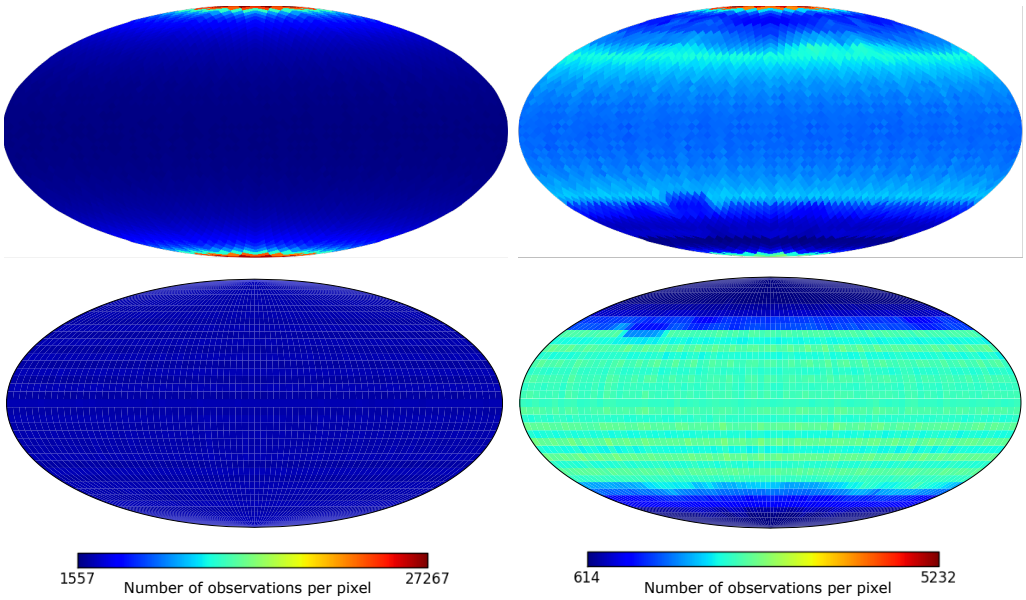

Figure 4.1: Plots showing the number of observations in each pixel for (top) HEALPix (Górski et al., 2005) and (bottom) equirectangular pixelizations, based on (left) all MONS prism-1 data and (right) only the frost-free data. There are approximately the same number of pixels in each subplot and the colorbars are shared within a column

et al., 2005) an equal area pixelization with quadrilateral pixels. Aside from this computational consideration it is the data that will drive pixelization choice. 2d histograms of the MONS data for these different pixelizations are shown in figure 4.1. These histograms have approximately the same number of pixels (3042 for equirectangular and 3072 for HEALPix), exact equivalence is not possible as HEALPix maps must have $12a^2$ pixels, with integer a. As can be seen in the plot, the HEALPix scheme leads to pixels at the pole having nearly 20 times the number of observations as those at the equator, though this discrepancy is less in the frost-free case due to much of the winter time data at the poles being removed. As the maximum entropy pixon reconstruction aims to find a representation of the image in which the information content is flat (section 2.5.1), we have selected the equirectangular pixelization as the most appropriate because it is closer to a flat information representation.

The results of figure 4.1 can be understood by considering the sampling of the MONS instrument. The Mars Odyssey satellite, from which the MONS data are measured, is

on an approximately polar orbit, with period 7070 s, and the instrument sampling rate is 19.2 s, which leads to approximately one sample per degree. The polar orbit means that the same time is spent in each latitude band. Equirectangular pixel area scales such that there are the same number of pixels in each latitude band, but as HEALPix pixels all have the same area, the number of pixels, in each latitude band, decreases towards the poles and the poleward pixels contain more observations. Also, the faint banding seen in the frost-free equirectangular plot (lower panels in figure 4.1) is due to beating between the instrument sample frequency and the pixel frequency in latitude (it is present in the non-frost-free case but less obvious due to the colour scheme).

The density of the pixelation should be chosen so that the reconstructed signal is well sampled, which will mean oversampling the PSF by several times. The final reconstruction in this chapter is pixelated into square, $0.5^\circ \times 0.5^\circ$, grid pixels in a equirectangular projection and the statistical errors within a pixel determined, empirically, by calculating the scatter between repeat observations in the observed count rate within the pixel. It was necessary to estimate the noise amplitude per pixel empirically as the various corrections applied to the data, for atmospheric thickness, cosmic ray flux, and spacecraft latitude and altitude, resulted in the data no longer possessing Poisson statistics (Maurice et al., 2011).

4.2.1 Assumed instrumental properties

As a starting point for our estimation of the MONS PSF we have taken the calculated prism-1 spatial response of the MONS instrument to a circular feature with 5 degree diameter from Prettyman et al. (2009). This calculation involved the creation of a detailed numerical model of the Martian surface and atmosphere and the Mars Odyssey spacecraft. The production of neutrons via cosmic ray interaction with the Martian atmosphere and surface and their subsequent transport within the surface, atmosphere and exosphere and then their interaction with the spacecraft and detection was performed using the Monte Carlo N-Particle eXtended (MCNPX) transport code (Pelowitz, 2005).

The result of this calculation is, effectively, a 'disk spread function'. To convert this to the instrumental PSF, we have calculated the result of convolving a 5° diameter disk with various trial PSFs, assuming that the PSF is azimuthally symmetric about the spacecraft's nadir point. The profile resulting from these convolutions that most closely matches that

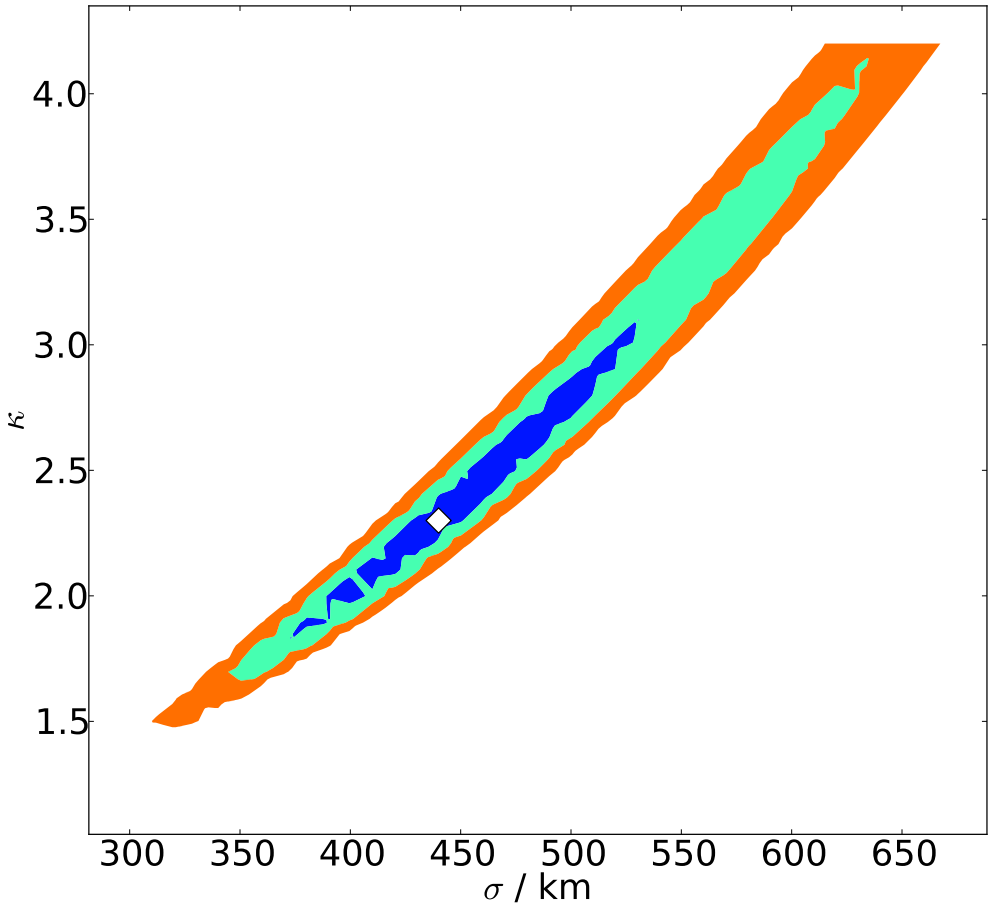

Figure 4.2: Constraints on the parameters in the PSF. The parameters are those in equation 4.2.1. The white diamond shows the best-fit reconstruction with minimum χ^2. Coloured regions enclose probabilities of 68%/95%/99.7%, determined using $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$.

obtained from the simulations is taken to be the correct PSF. We make the assumption that the PSF can be described by a kappa function, defined as

$$\frac{f(s)}{f(0)} = \left(1 + \frac{s^2}{\sigma^2}\right)^{-\kappa}, \tag{4.2.1}$$

where κ and σ are parameters to be determined, s is the arc length from the sub-orbital point and $f(s)$ is the response of the instrument to a source at s. κ functions are often used to model the response of spacecraft-borne gamma ray and neutron detectors (Maurice et al., 2004). The results of this calculation are shown in Fig. 4.2, from which the best fitting PSF has $\sigma = 440$ km and $\kappa = 2.3$, giving a FWHM of 520km.

Table 4.1: Parameters used in epithermal neutron to wt. % WEH conversion.

Parameter	Value
a_0	0.6991
a_1	0.3291
a_2	-0.1152
a_3	0.01986
a_4	-0.001651
a_5	5.12×10^{-5}

4.2.2 Conversion of epithermal neutron flux to water equivalent hydrogen measurements

All of the conversions from epithermal neutron count rate to weight % water equivalent hydrogen (wt. % WEH) in this thesis use the single layer model in Feldman et al. (2004), i.e. the surface is assumed to be made up of a semi-infinite layer with a constant wt. % WEH content. The conversion is a polynomial fit relating detected epithermal neutron count rate to wt. % WEH based on the result of MCNPX simulations. Specifically, the wt. % WEH content, M_{H_2O}, is related to epithermal neutron flux, c_{epi}, via

$$M_{H_2O} = \sum_{i=0}^{5} a_i c_{epi}^i, \tag{4.2.2}$$

where the constants fitted by Feldman et al. (2004) a_i, are defined in table 4.1. The function, along with the limits of applicability, are plotted in figure 4.3.

This scaling will be a poor approximation over much of the surface which is covered with a layer of relatively dry dust and also in regions that show lateral inhomogeneity. The conversion is included here simply as an illustration and if the absolute value of the WEH abundance at a particular location is needed then a location specific model should be used.

4.3 Results

The global frost-free MONS data and our reconstruction of the underlying field are shown in the two panels of figure 4.4 Locally adaptive pixon reconstruction of the data leads to

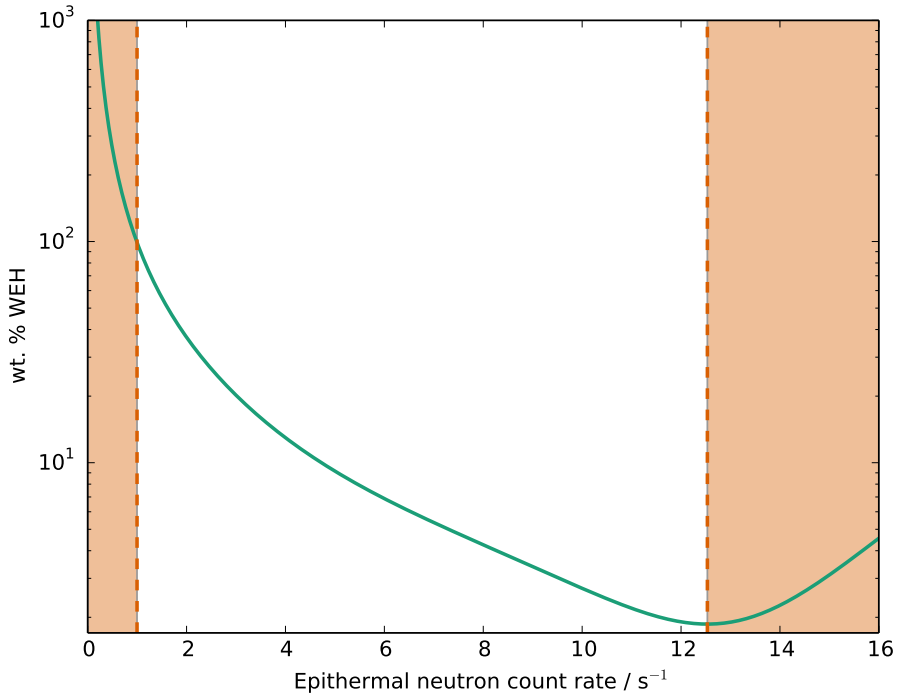

Figure 4.3: Polynomial fit relating observed epithermal neutron count rate to wt. % WEH content. The red shaded regions denote count rates for which the conversion should not be believed, at less than $\sim$1 count per second the fit predicts more than 100 wt. % WEH and at count rates greater than $\sim$12.5 counts per second the WEH abundance begins to rise with count rate. Both of these regimes are unphysical.

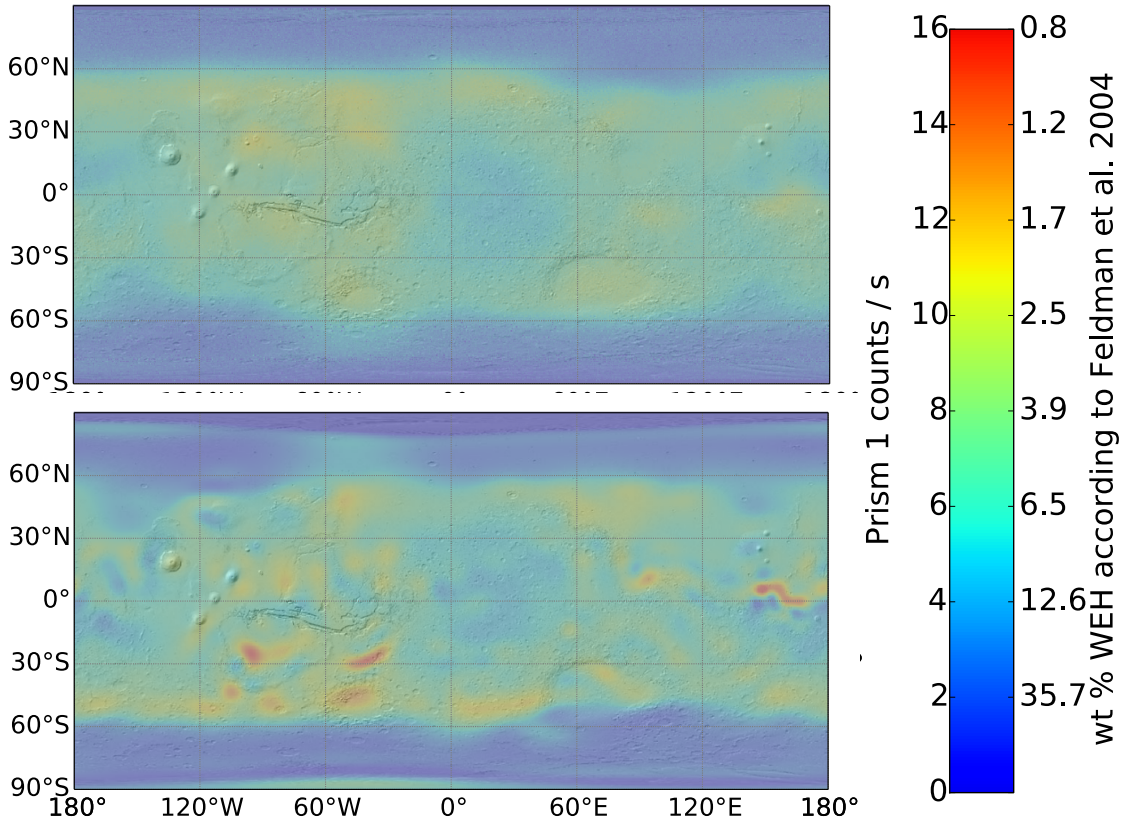

Figure 4.4: *Top.* The global frost-free MONS prism-1 data, used in the reconstruction. *Bottom.* A global locally-adaptive pixon reconstruction of the MONS prism-1 data. Underlayed on both panels is a MOLA shaded relief map. The colour bar also shows the conversion to wt. % water equivalent hydrogen from Feldman et al. (2004).

an increase in dynamic range of the count rate of nearly 50% is seen, from ∼1-11 to ∼0-16 counts per second. In addition to improvement in dynamic range there is and enhancement in spatial resolution, quantified below.

4.3.1 Improvement in resolution

To estimate the resolution of the reconstructed data we consider the two-point correlation function of the reconstruction. Comparing the separation required for the correlation function to drop to a given level allows the effective resolution of the reconstruction to be calculated. The correlation between the count rate at two points (with angular seperation θ) on the surface of the planet can be calculated using the relation

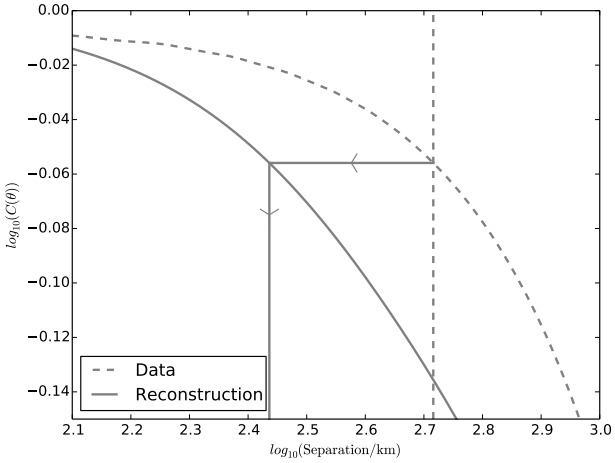

Figure 4.5: The two point angular correlation function of the MONS data and the reconstruction, normalized to unity at zero separation. The vertical dashed line shows the MONS PSF FWHM scale. The solid vertical line at 290 km shows the separation required for the power in the reconstruction to drop by the same proportion that the power in the data drops at a separation equivalent to the FWHM of the PSF.

$$C(\theta) = \frac{1}{4\pi}\sum_{l=0}^{\infty}(2l+1)C_l P_l \cos\theta \tag{4.3.3}$$

where P_l are the Legendre polynomials and C_l is the power spectrum of the count rate, defined by

$$C_l = \frac{1}{2l+1}\sum_{m=-l}^{l}|a_{lm}|^2 \tag{4.3.4}$$

and a_{lm} are the spherical harmonic coefficients defining the count rate field.

This function, for the data and reconstruction, is plotted in figure 4.5 and shows that the effective FWHM of the map decreases from 520 km to 290 km, i.e. variations on scales smaller than 290 km are unlikely to be found in our reconstruction. This three fold decrease in effective detector footprint leads to the enhanced dynamic range shown in the reconstruction in figure 4.4. The value of $C(0)$ i.e. the power on the smallest scales, increases by 40% between the data and reconstruction.

4.3.2 Estimation of errors in the pixon method

To estimate the systematic and random errors in the pixon method, when applied to the MONS data, we created a set of 10 mock data sets. These mock data were generated by blurring the reconstructed image (figure 4.4 lower panel) with the PSF and then adding uncorrelated Gaussian white noise with σ as determined for the real data. Running the pixon reconstruction algorithm on each of these mock data sets enables the errors in the reconstructed count rate, both statistical, *s*, and systematic, to be estimated. The upper panel of Figure 4.6 shows the RMS scatter in each pixel over the 10 reconstructions, a measure of the statistical error in the reconstruction due to noise in the data. Comparison with figure 4.4 shows the noise to be greatest where the count rate is large or varies quickly, such as at the south polar cap. The global average error in the reconstructed epithermal count rate is 0.26 counts per second. The average difference between a reconstruction based on mock data from one based on the real data is shown in Figure 4.6 (lower panel). This provides a measure of the systematic errors in the reconstruction method and has an average of 0.47 counts per second. The regions with highest offset are, mostly, colocated with those with the largest random errors. The feature at around 0° N, 150° E is particularly discrepant as it features regions of very low and high count rate in close proximity. The inability to completely remove the smoothing effect of convolving with the PSF causes these regions to blend together. The mottling in these images is an artefact related to the size of the pixons.

4.4 Comparison of image reconstruction techniques

Several versions of the pixon method, and other popular image reconstruction techniques, were described in chapter 2; in this section these techniques will be applied to a mock data set similar to the MONS data and the most effective technique determined. The mock data are based on a locally adaptive pixon reconstruction of the MONS data which has been convolved with the MONS PSF, as determined in section 4.2.1, and random gaussian noise, with the statistics determined in section 4.2, added. As in chapter 2, to

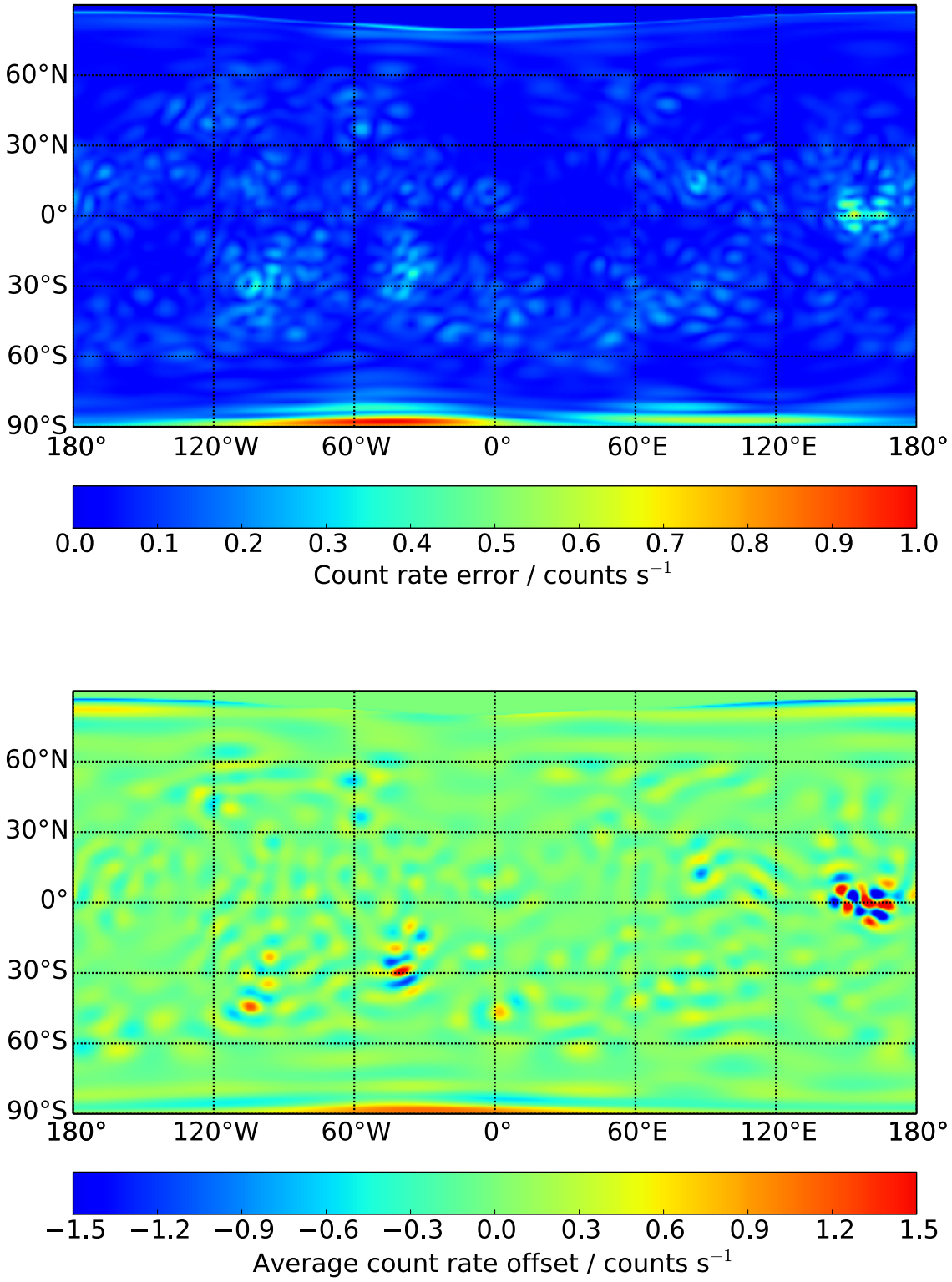

Figure 4.6: *Top.* The standard deviation of the reconstructed count rate based on 10 mock data sets. *Bottom.* A plot of the average difference, in each pixel, between the mock reconstructions and the true image.

compare these reconstructions we define (by analogy with χ^2) the misfit statistic ψ^2,

$$\psi^2 = \sum_{\mathbf{x}} \frac{(I(\mathbf{x}) - \hat{I}(\mathbf{x}))^2}{s^2}, \tag{4.4.5}$$

where s is that statistical error in the reconstructions (figure 4.6 upper panel). The values of this statistic for the different reconstructions are shown in figure 4.7. The figure shows that the Bayesian pixon methods give rise to a more accurate reconstruction that the iterative Jansson's method. The order of magnitude variation in E_R in the Jansson's method reconstructions is simply a statistical fluctuation, the probability distribution has a long flat tail, and does not suggest that a statistically poorer fit was achieved in one of the reconstructions. Of all the pixon methods the locally adaptive version is seen to be the most suitable for reconstruction MONS-like data.

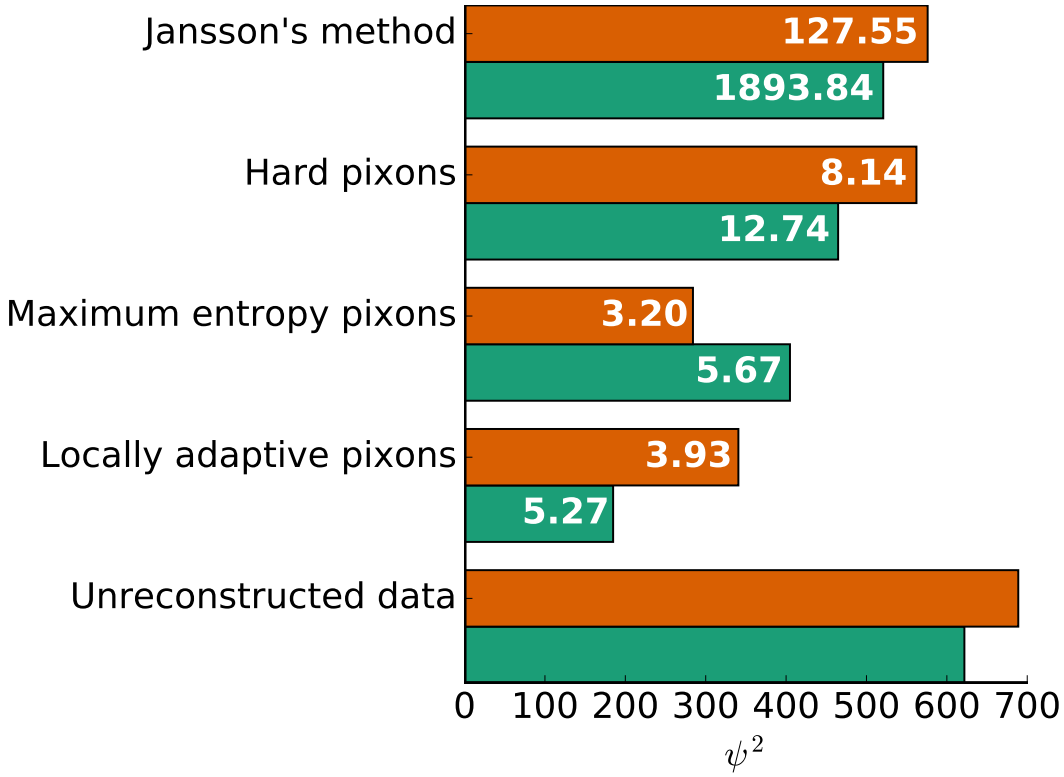

Figure 4.7: The ψ^2 misfit statistic calculated for several reconstruction techniques, labelled on the left, and taking the mock data itself to be the inferred truth, $\hat{I}$. The numbers in the bars indicate the values of the misfit statistic, E_R, lower numbers correspond to a closer fit to the true image. The red and green bars correspond to mocks based on maximum entropy and locally adaptive pixon reconstruction of the data, respectively.

Figure 4.10 shows the results of these reconstructions along with the image and data on which they are based.

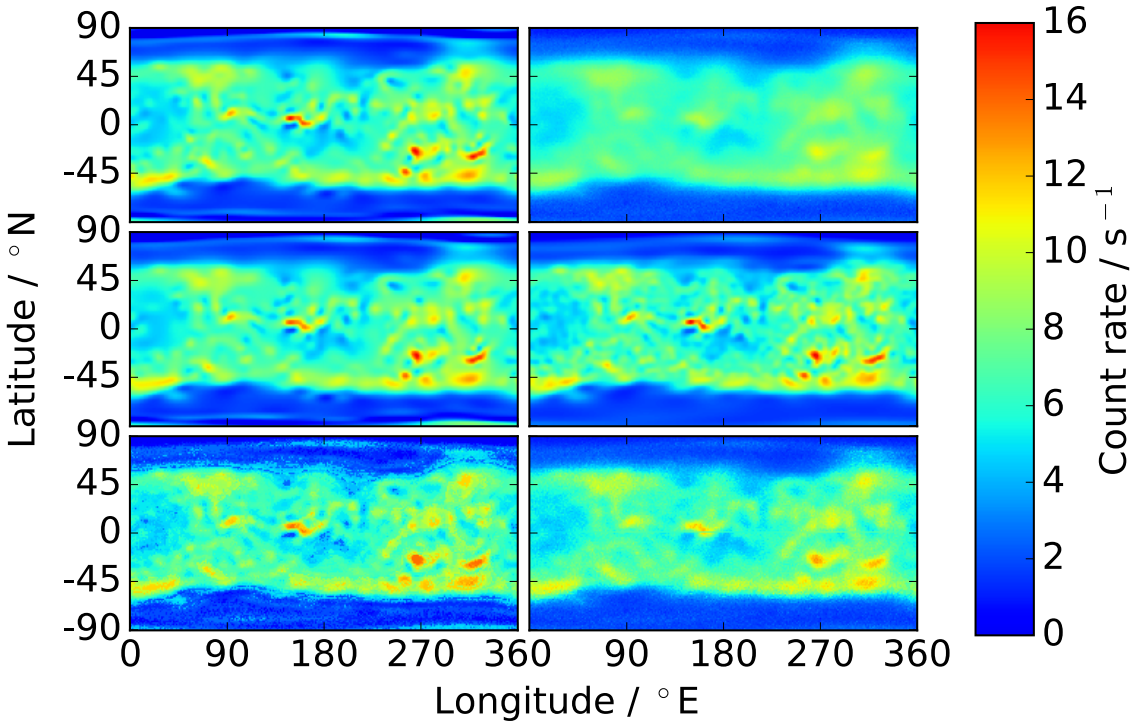

Figure 4.8: Cylindrical projections showing, clockwise from top left, the true image on which the mock data are based, the mock data i.e. the blurred and noisy image, the maximum entropy pixon, Jansson's, hard pixon and locally adaptive pixon reconstructions of the mock data.

The conventional, maximum entropy pixon method fails to provide sufficient freedom to reconstruct features where the signal to noise ratio in the image is low, such as at the poles (see figure 4.9), and has too large pixons where the image has a high signal to noise ratio but no spatial variation, e.g. in the Terra Sabaea centred at 2° N, 42° E. It should be noted that although the hard pixon reconstruction is less accurate than the other pixon methods it has a significantly larger posterior probability. This fact is a result of the small number of pixons required to describe the reconstruction and the neglecting of pixon structure, i.e. which pixels are in which pixons, when calculating the prior. Altering the prior so that this complexity is taken into account, or enforcing the contiguity of pixons should help improve the method, but we will leave this for future work and focus on the locally adaptive pixon method for the remainder of this thesis.

4.5 Comparing reconstructions with surface measurements

Our reconstructions also enable us to examine the region around Gale crater, the landing site and science target of the Curiosity rover. The Curiosity rover has an active neutron sensor onboard, the Dynamical Albedo of Neutrons (DAN) instrument (Mitrofanov et al., 2012). Unlike for passive, orbital remote sensing, DAN works by illuminating the surface with a pulsed neutron source and then measuring the returned neutron flux to infer composition. This gives us a ground truth against which to compare our reconstructions.

The inferred hydrogen abundance at Gale crater changes little between the raw data and the reconstruction (lower panels in Figure 4.11), being around 8.0 ± 0.4 wt.% WEH in both cases. However we find a gradient in the higher resolution reconstruction, with hydration increasing towards the south-east. It is worth noting that these results are discrepant with those obtained by DAN, which suggest an average abundance of 2.1 to 2.7 wt% WEH (Mitrofanov et al., 2014) and those from the Mars Odyssey Gamma Ray Spectrometer suite with 5 wt% WEH (Mitrofanov et al., 2014). It has been suggested that this discrepancy may be the result of the differing sizes of the response functions of the ground and space based detectors (3 m and 520 km respectively) and the presence of the rover in a locally dry region (Mitrofanov et al., 2014). However, if this were the correct explanation we may expect that improving the resolution would alleviate this discrepancy and we find no evidence of this in our reconstruction. For this explanation to hold, the variation in abundance must be on a scale smaller than the scales accessible to the reconstruction (i.e. 290 km). Alternatively, the differing mean depths from which the detected neutrons originate may be responsible: DAN is sensitive only to the composition of the top 60 cm of the surface (Mitrofanov et al., 2014), whereas the orbital detectors see up to 1 m into the surface, depending on the soil water abundance. If the deeper regions are more hydrated than the surface then we would expect the orbital detectors to measure larger WEH values than DAN.

4.6 Conclusions

We compared the performance of several reconstruction techniques to simulated data sets, designed to resemble the MONS data, and found the maximum entropy and locally adap-

tive pixon methods to be the most effective. The difference in performance between the two techniques is small but the locally adaptive pixon method is favoured. Using our new locally adaptive pixon reconstruction technique we have produced the highest resolution map of the near-subsurface hydrogen abundance on Mars (figure 4.12) available to date. This map has an effective resolution of 290 km, a near-two-fold improvement over previously available data sets. Along with the increase in resolution the dynamic range of epithermal neutron count rate is increased by 50%, revealing greater hydrogen contents than were seen in the unreconstructed data. We will examine some of the implications of these newly revealed hydrated features in the next chapter.

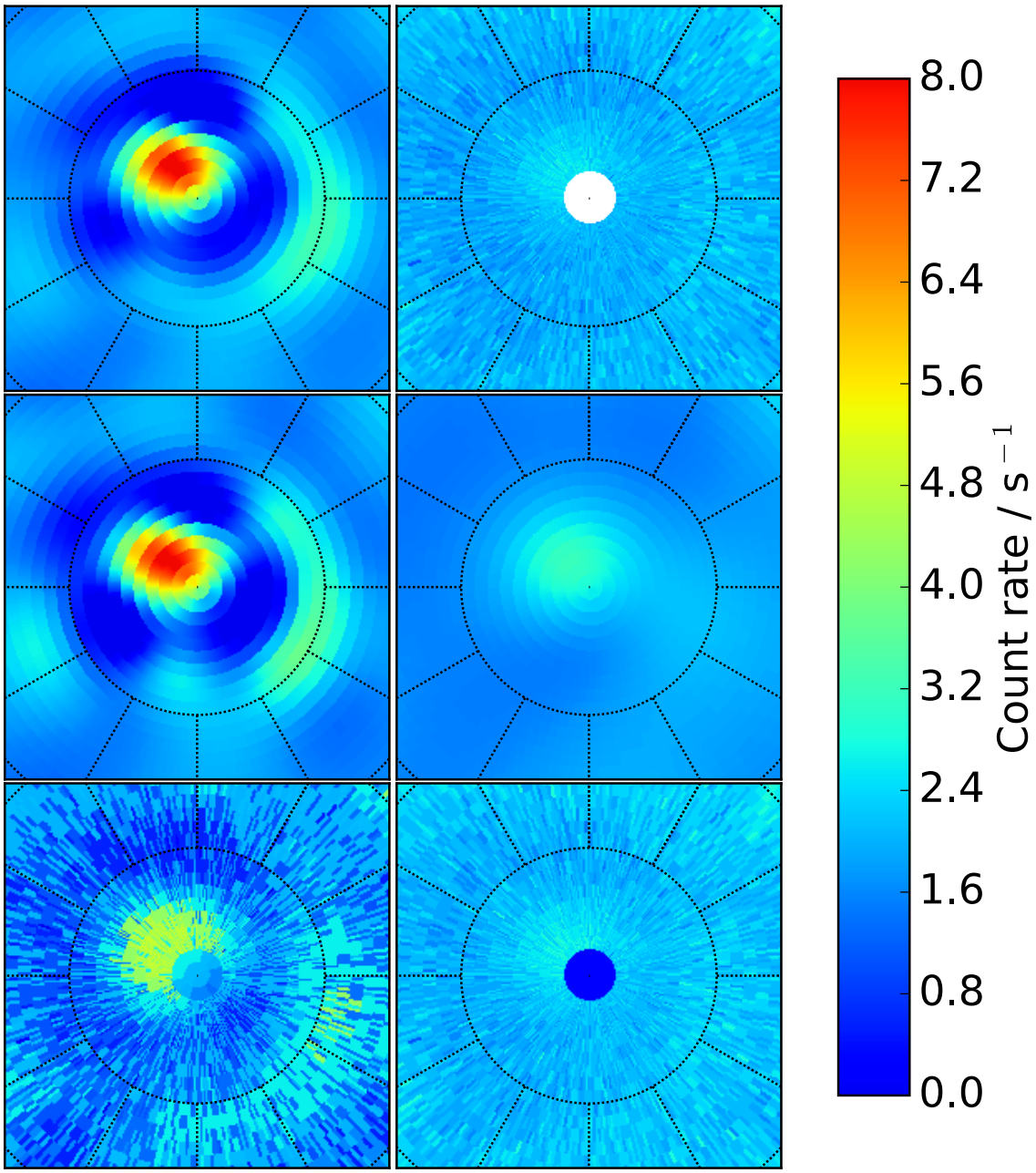

Figure 4.9: Stereographic projections, centred on the south pole and going out to 75°S, showing the mock data and various reconstructions. Clockwise from top left the true image on which the mock data are based, the mock data i.e. the blurred and noisy image, the maximum entropy pixon, Jansson's, hard pixon and locally adaptive pixon reconstructions of the mock data. The true image is a locally adaptive pixon reconstruction of the frost-free, MONS data.

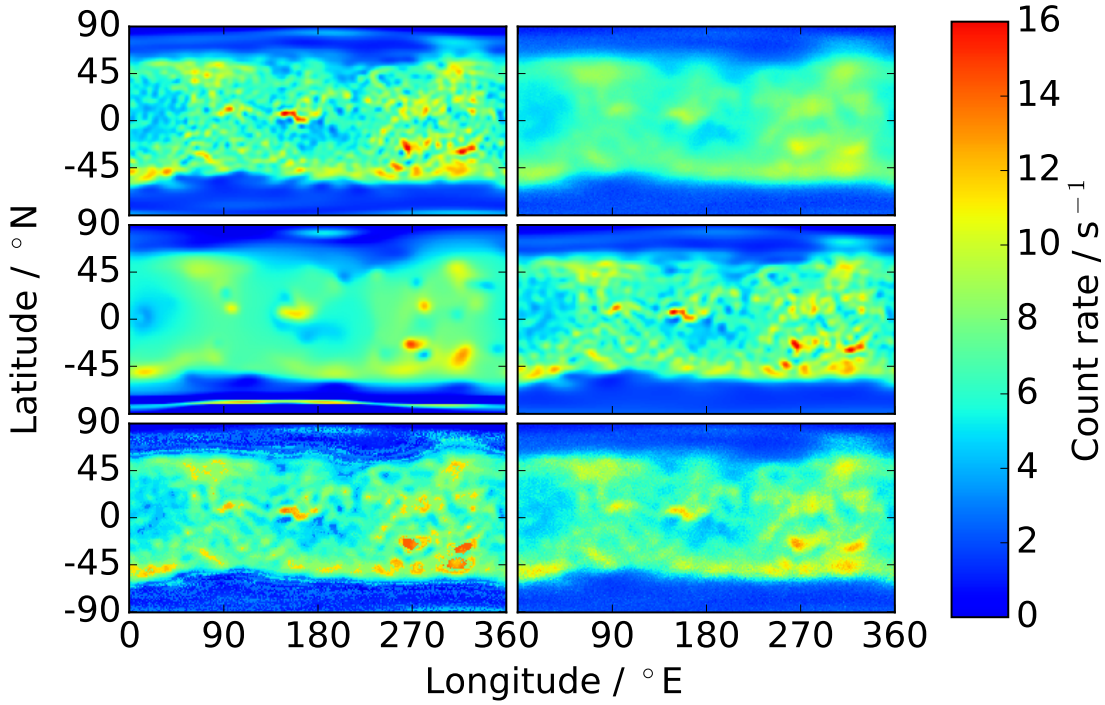

Figure 4.10: Cylindrical projections showing, clockwise from top left, the true image on which the mock data are based, the mock data i.e. the blurred and noisy image, the maximum entropy pixon, Jansson's, hard pixon and locally adaptive pixon reconstructions of the mock data. The true image is a maximum entropy pixon reconstruction of the frost-free, MONS data.

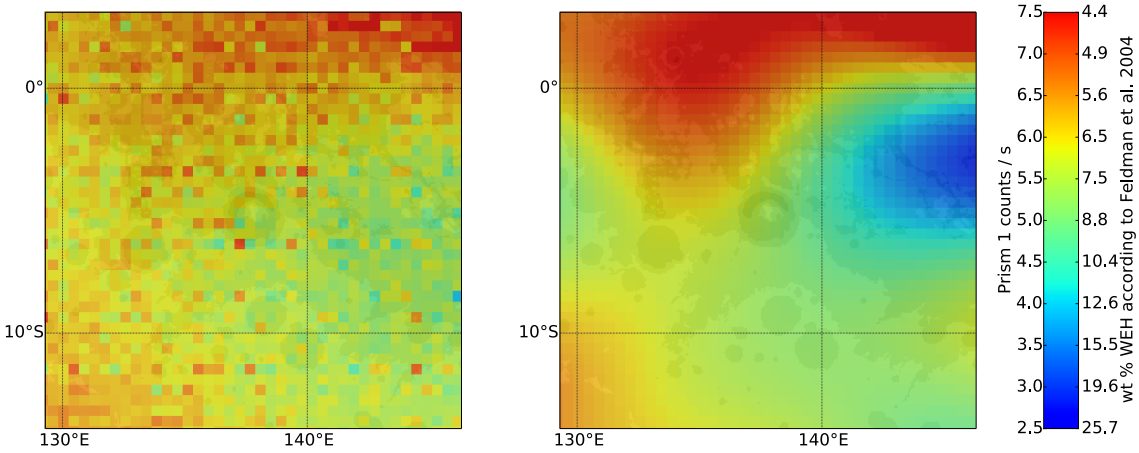

Figure 4.11: The MONS data (left) and reconstruction (right) in the area around Gale Crater. Underlayed is a MOLA shaded relief map. The colour bars also show the conversion to wt. % water equivalent hydrogen from Feldman et al. (2004).

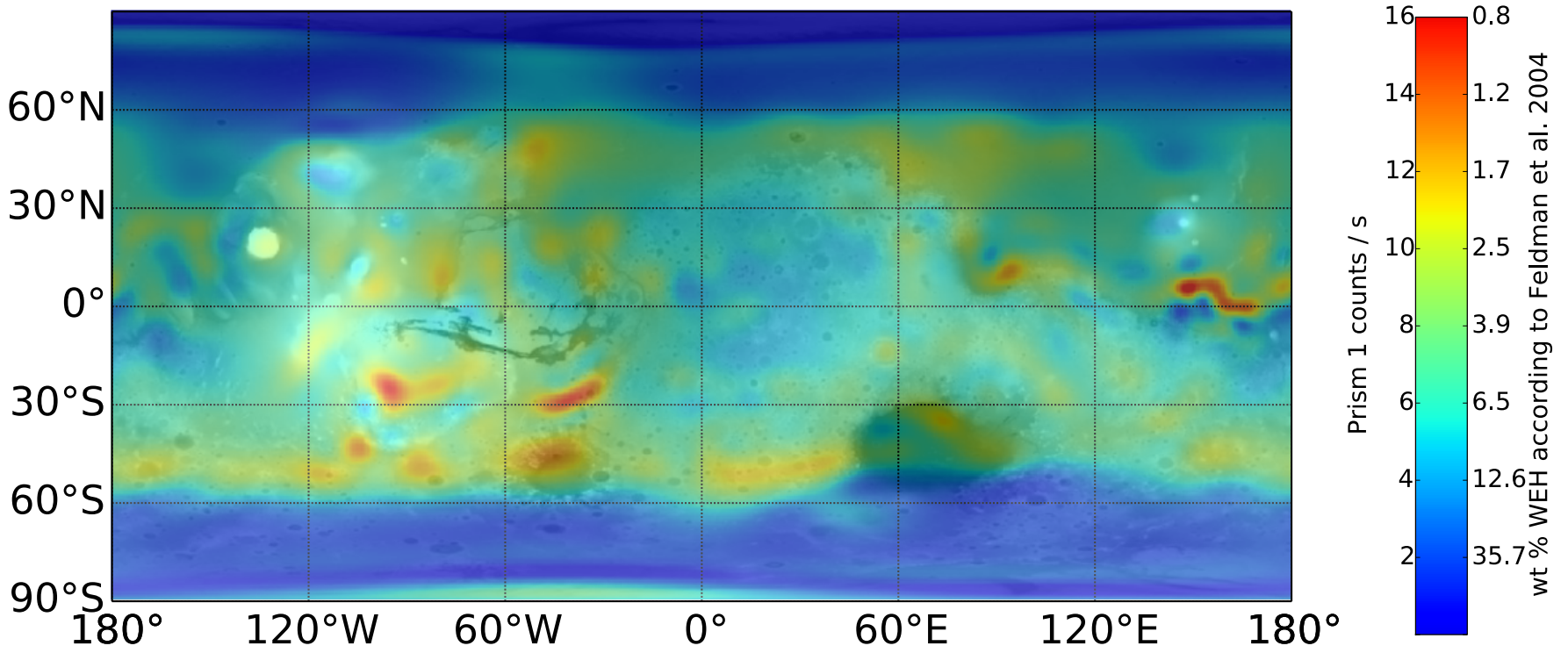

Figure 4.12: Our final locally adaptive pixon reconstruction of the MONS data. Underlayed is a brightness map corresponding to MOLA topography. The colour bar shows the conversion to wt. % water equivalent hydrogen from Feldman et al. (2004).

CHAPTER 5

Small scale features in the MONS reconstruction

We showed, in the previous chapter, that pixon reconstruction of the MONS data yields a near doubling in spatial resolution and a 50% increase in dynamic range. Here, we will use these results to examine a series of science targets that have sizes close to or smaller than the MONS PSF and are suggested, or believed, to contain contemporary water ice or hydrated mineral. The selected sites include the north pole, the sites of Recurring Slope Lineae activity in the southern mid-latitudes and some equatorial regions outlined in figure 5.1.

Using the improved resolution map of the epithermal neutron count rate across the entire surface of Mars, from the previous chapter, we find evidence for large amounts of buried, non-polar hydrogen in the form of both hydrated minerals (at Meridiani Planum) and possible buried water ice (in the Medusae Fossae Formation and on the slopes of the Tharsis Montes). We find that the reconstructed data are able to distinguish water sources larger than 120 km in diameter. The increase in resolution over the unreconstructed MONS data allows better correspondence structures in the neutron data and features in surface imagery, thus enabling a more robust geophysical interpretation.

5.1 Elysium Planitia: ice or lava?

The Cerberus fossae are a 1600 km long set of parallel fissures at Elysium planitia, in equatorial Mars. The fissures stretch from 16.16° N, 154.43° E to 6.23° N, 174.72° E. They are shown in a Mars Orbiter Laser Altimeter, MOLA, topographic image in figure 5.2. The fault is believed to be related to the Elysium Montes, located to the northwest,

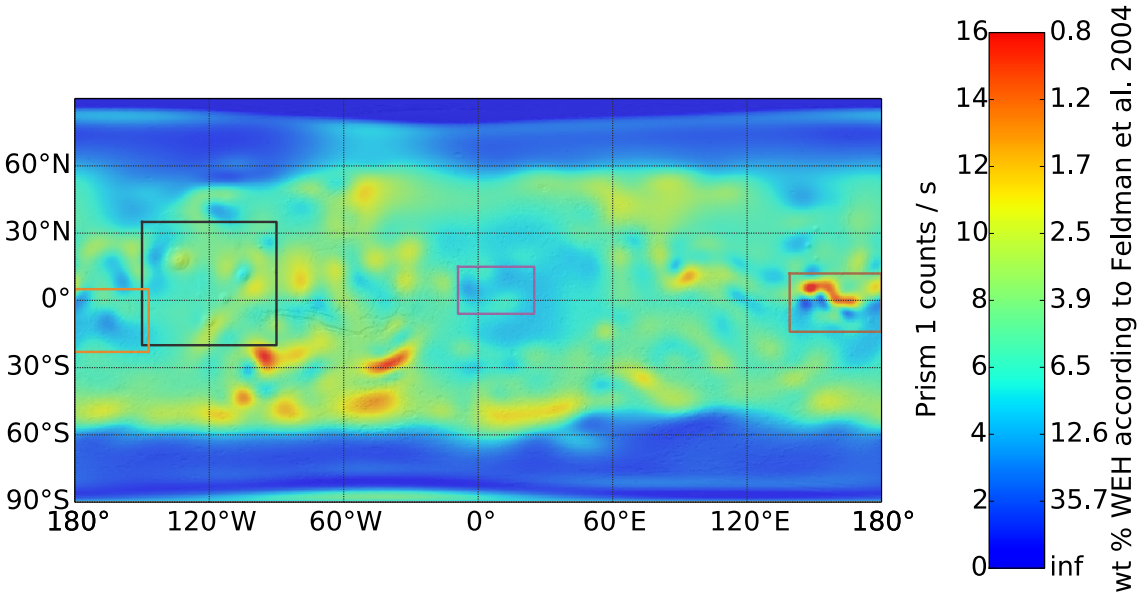

Figure 5.1: Global MONS reconstruction outlining the regions examined in later sections. The Tharsis Montes, examined in section 5.2.2, are outlined in black, Elysium planitia and part of the Medusae Fossae Formation, in brown, the remaining section of the Medusae Fossae Formation, in orange, and Meridiani Planum in violet.

and there is morphological evidence that it has been the source of both water and lava floods (Berman & Hartmann, 2002; Burr et al., 2002; Head & Marchant, 2003) within the last 2-10 Myr. It had been assumed that the water had sublimated away, leaving only the fluvial channels of Athabasca and Marte Valles as evidence of the water flows. However, Murray et al. (2005) identified, in Mars Express High Resolution Stereo Camera (HRSC) images, plate-like features that they argued must be the remains of a buried frozen sea 800×900 km across and up to 45 m deep, centred at 5° N, 150° E. Murray et al. (2005)'s argument is based on noting the geomorphological similarities of the HRSC images with those taken above Antartica showing the break up of pack-ice, discrepancies between the ages of the plates and inter-plate regions, along with the reduction in the volume of craters that have been filled by the flow. Part of figure 1 of Murray et al. (2005) is reproduced here in figure 5.3, which shows the similarity in morphology between the Elysium plate-like features and Antarctic pack-ice.

The epithermal neutron flux is sensitive only to the top few tens of centimetres of the surface, which, if the frozen sea were present, may consist of sublimation lag or pyroclastic deposits from subsequent eruptions from the Cerberus fossae. Nonetheless, if

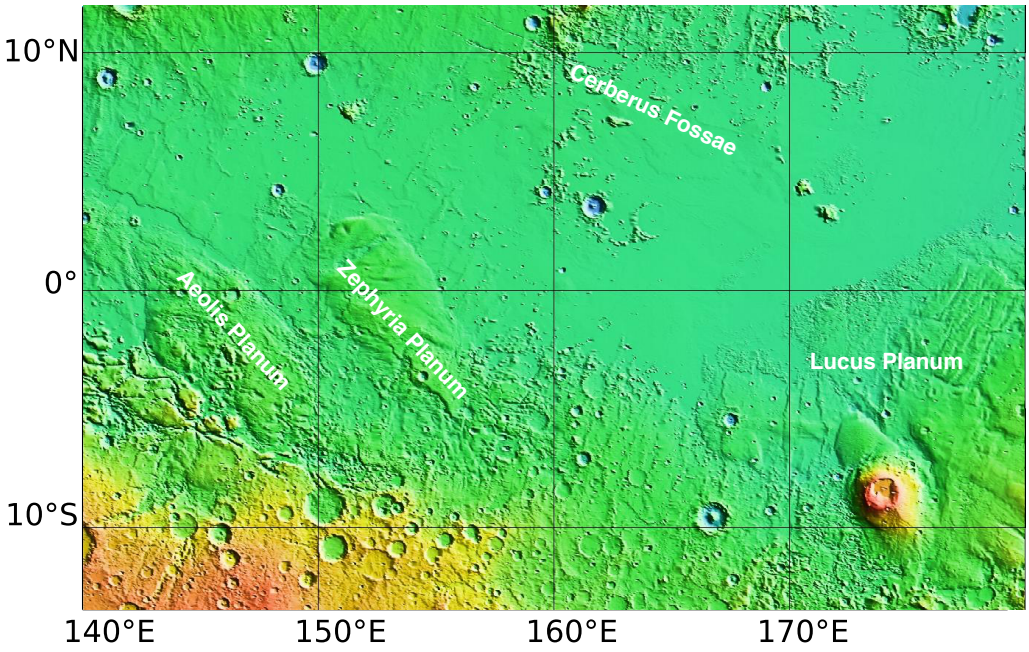

Figure 5.2: A MOLA topographic image of the Southern part of Elysium Planitia, identified by Murray et al. (2005) as containing a large deposit of buried water ice. The eastern part of Cerberus fossae are visible at the top centre, other parts of the fossae are beneath the resolution of this image. The lobes of the Medusae Fossae Formation are also labelled.

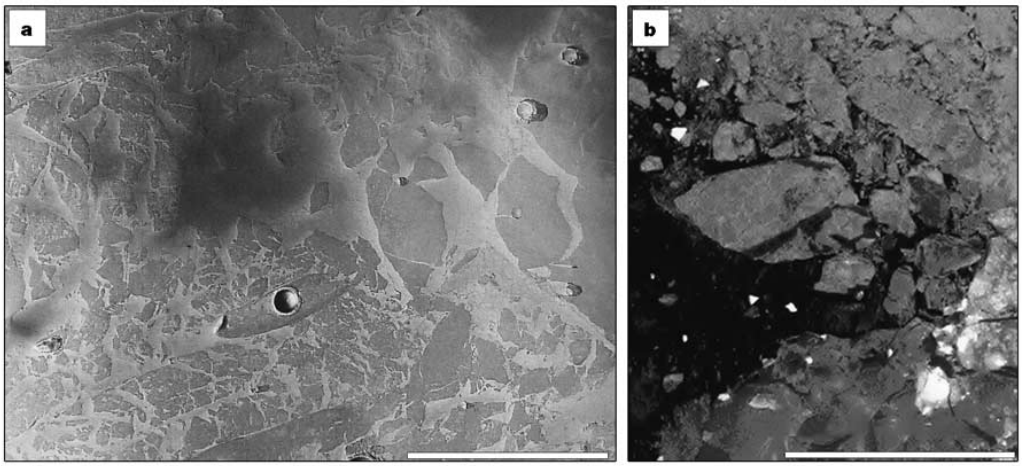

Figure 5.3: **a**, Section of a HRSC image centred at 5.5° N, 150.4° E, showing plate-like features. **b**, Synthetic Aperture Radar image of pack-ice (credit:ESA). The scale bars in both panels are 25 km.

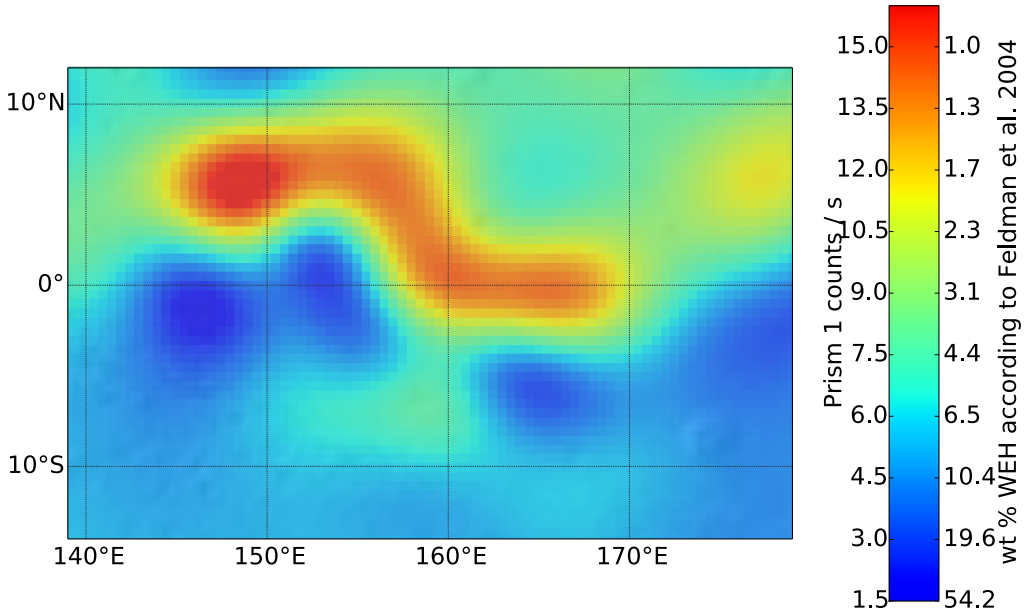

Figure 5.4: A zoom of the locally adaptive pixon reconstruction on the same region as figure 5.2.

the outflow occurred within the last 10 Myr we may still expect to see excess hydration due to partial protection of the water by a dust cover lag, perhaps containing hydrated mineral layers that provide a barrier protecting a deeper, higher-humidity environment (Feldman et al., 2011; Wang & Ling, 2011).

The locally adaptive pixon reconstruction of the region around the proposed, buried sea is shown in figure 5.4. The location of the water ice sea, 5 N and 150 E, corresponds to one of the locations with the highest epithermal neutron flux on the entire surface of Mars. This suggests that the top few tens of centimetres of soil, in this region, are unusually dry with < 1 wt. % WEH. The dry feature in the reconstruction extends beyond the region identified by Murray et al. (2005) and covers much of the smooth plains that are believed to be young basaltic lavas from Cerberus fossae. Previous work using MONS data has considered only the effects of H and CO_2 on the epithermal neutron flux, to understand the effect of changing the abundances of other elements, such as iron, will require new neutron transport simulations.

The results of section 4.3.2 indicate a systematic error of up to 1.5 counts per second and a random error of 0.6 counts per second in this region. These errors are amongst the highest on the surface, however even in the worst case scenario that the random and

systematic errors were in the same direction southern Elysium planitia remains an exceptionally dry region. Thus, our conclusions will not be affected, qualitatively, by errors in the reconstruction technique.

5.2 Possible sites of equatorial hydration

In this section we will focus on a few locations that have been proposed to contain water in the equatorial regions of Mars deposited during periods of high orbital obliquity. Again, we will use the locally adaptive pixon reconstruction of the MONS data to assess hydrogen (and therefore hydration) abundance in the near-subsurface Martian soils. Evidence for equatorial hydration is both morphological (Head & Weiss, 2014) and compositional (Feldman et al., 2004). This water is hypothesised to have been deposited during past periods of high orbital obliquity when the water ice currently at the poles becomes unstable and is ultimately deposited elsewhere (Forget et al., 2006).

5.2.1 The Medusae Fossae Formation: ice or dry, porous rock?

The Medusae Fossae Formation (MFF) is a discontinuous geological unit of easily erodible material that stretches $\sim$ 1000 km across equatorial latitudes, along the boundary of the northern lowlands and southern highlands, located in both Elysium and Amazonis Planitiae. It consists of five main lobes: the Aeolis and Zephyria Plana in the west (figure 5.2), Lucus Planum at the centre, and Eumenides Dorsum and the combined Amazonis Mensae and Gordii Dorsum in the east (figure 5.5), though many smaller outliers are identified in the southern highlands (Harrison et al., 2010). The MFF consist of three stratigraphic members, distinguished by age, morphology, albedo and stratigraphic relations (shown in the geological map in upper panel of figure 5.5 as distinct units, see (Tanaka et al., 2014) for an updated map).

The origin of the MFF is uncertain. Proposed explanations include consolidated pyroclastic deposits (Scott & Tanaka, 1982) and aeolian sediments with ice-rich material (Head & Kreslavsky, 2004), similar to that found in the polar layered deposits (Schultz & Lutz, 1988), laid down during periods of high orbital obliquity. Radar sounding using the MARSIS instrument onboard ESA's Mars Express has been used to measure the dielec-

tric constant of the MFF material and found it to be consistent with the MFF containing a large component of water ice (or anomalously low density soil) (Watters et al., 2007). It may be possible to distinguish between these two mechanisms using neutron derived hydrogen abundances.

The topography and MONS reconstruction of the western lobes of the MFF are shown in figures 5.2 and 5.4, and for the eastern lobes in figure 5.5. The western lobes, Aeolis, Zephyria and Lucus Plana, of the MFF are greatly enriched in hydrogen, with $>$10 wt. % WEH. At Aeolis Planum the neutron data, taking the errors from section 4.3.2 into account, imply a WEH abundance $>$ 30 wt. %. However, the eastern lobes of the MFF contain little, if any, increased hydration. Geologically, the western lobes are associated with the lower members of the deposit, whereas the eastern lobes corresponds to upper and middle members only (Scott & Tanaka, 1986; Greeley & Guest, 1987). The region of enhanced hydration at 10° S, 165° W is not coincident with the MFF, but does lie on the dichotomy boundary.

This result is consistent with the lower member of the MFF containing ice rich material, which lends weight to the theory that, at least part of, the MFF is a polar layered-like deposit. However, (Clifford et al., 2011) suggest that, even if the MFF deposits were initially dry, they may have become charged with ice in a geologically short span of time if the underlying unit were hydrated, so this result alone does not rule out other origins of the feature.

5.2.2 Tropical mountain glaciers at the Tharsis Montes

Evidence, in the form of surface morphology and cratering, for late Amazonian tropical mountain glaciers on the north-western flanks of the Tharsis Montes and Olympus Mons has been detailed extensively using observations from the Mars Express, Mars Global Surveyor and Mars Odyssey orbiters (Head & Marchant, 2003; Shean et al., 2005, 2007). The extent of these features is illustrated with white outlines in the right panel of figure 5.6. The production of such glaciers at equatorial Mars today is impossible given the current climatic conditions. Thus, the existence of these glaciers is interpreted to be the result of the migration of volatiles (chiefly water) from the poles to the equator during past periods of high orbital obliquity. Climate models predict the accumulation of ice on the

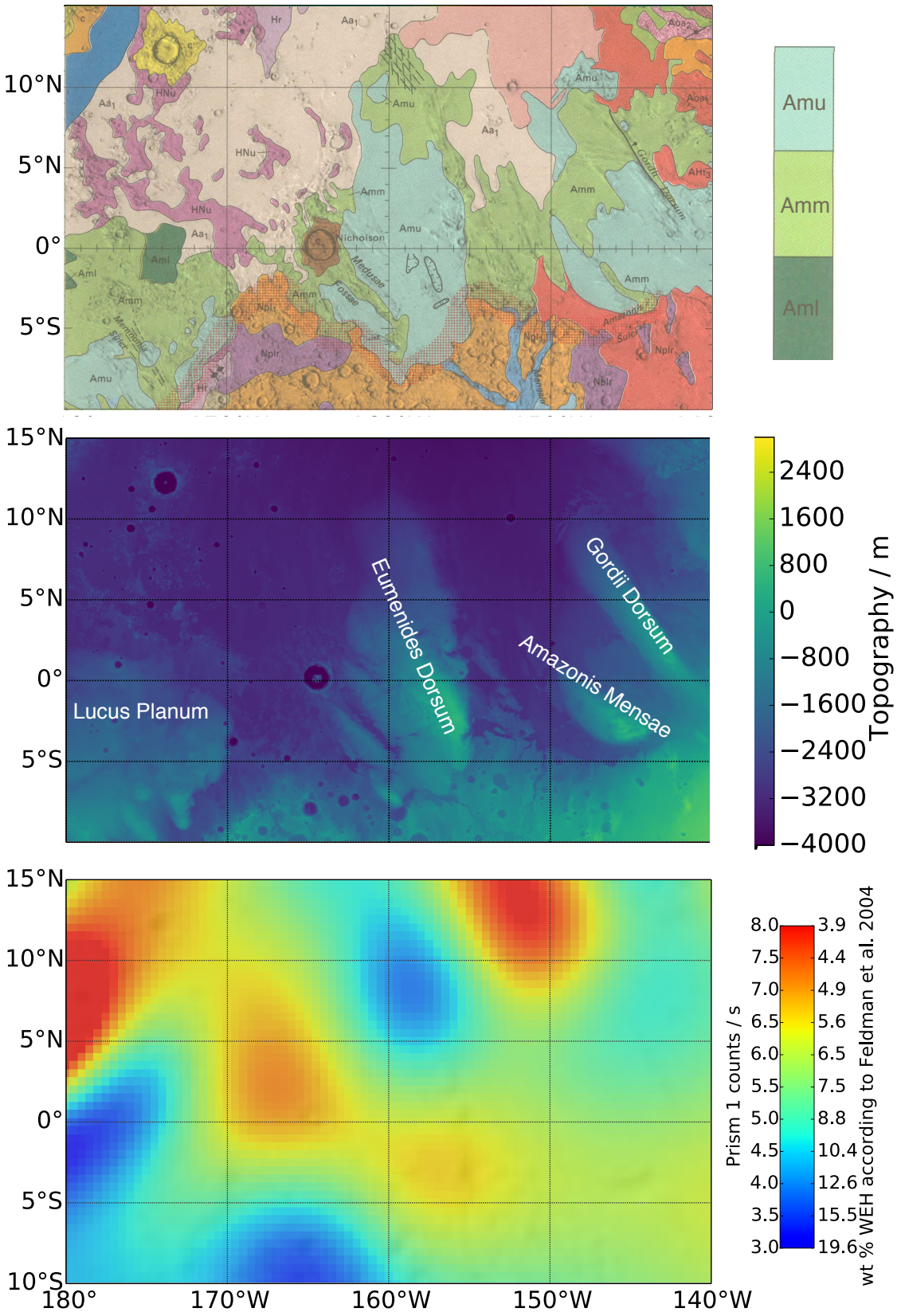

Figure 5.5: The eastern lobes of the Medusae Fossae Formation as (*Top*) a geological map, adapted from Greeley & Guest (1987), Amu, Amm and Aml are the upper, middle and lower members of the MFF, respectively, (*Centre*) MOLA topography, including lobe labels, and (*Bottom*) locally adaptive pixon reconstruction of MONS epithermal neutron data.

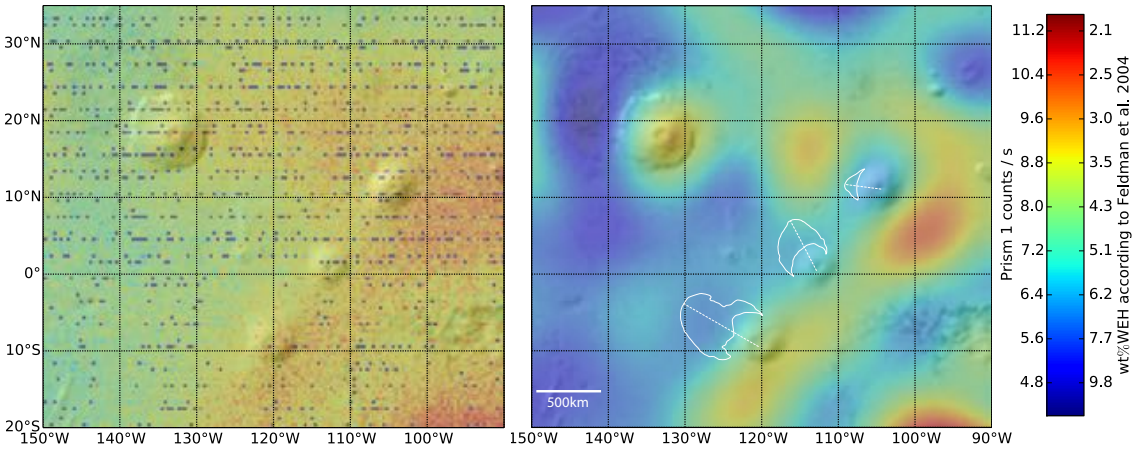

Figure 5.6: **Left**: Gridded MONS prism-1 data at Tharsis. **Right**: Pixon reconstruction of the MONS prism-1 data at Tharsis, showing the WEH conversion of Feldman et al. (2004). Underlaid is a MOLA shaded relief map. The white contours show fan shaped deposits on the slopes of the Tharsis Montes (and are based on those in Fastook et al. (2008).

north-western slopes of these mountains, during hypothesised periods of high obliquity, due to the adiabatic cooling of moist polar air (Forget et al., 2006; Madeleine et al., 2009).

The extent to which these equatorial deposits remain and the form in which they are present is not yet a settled question. Campbell et al. (2013) see no evidence for buried ice at Pavonis Mons using the Shallow Radar (SHARAD) instrument onboard the Mars Reconnaissance Orbiter. However, morphological evidence is presented by Head & Weiss (2014) in the form of fresh ring-mold craters at both Pavonis and Arsia Mons, which is suggestive of the presence of buried ice in the, geologically, very recent past.

We have used the pixon image reconstruction method on data from the Mars Odyssey Neutron Spectrometer (MONS) to produce a high resolution map of the Water Equivalent Hydrogen (WEH) distribution across the surface of Mars. Despite the fact that the present climate prevents the stable presence of subsurface ice within the range accessible to the neutron data ($\sim$1 m of the surface) Elphic et al. (2004) showed, also using MONS data, that elevated hydrogen contents are seen around the slopes of the Tharsis Montes. However the utility of this data set in constraining the distribution of water on the scale of the tropical mountain glaciers has been limited by its poor spatial resolution due to the 520 km FWHM footprint of the MONS.

Jansson's method was used by Elphic et al. (2005a) to perform image reconstruction

on the MONS data at Tharsis. However, it is known that this method amplifies noise and may introduce spurious features (Prettyman et al., 2009). Here we have used the locally adaptive pixon image reconstruction technique to improve the spatial resolution of this analysis by suppressing noise and removing the effect of blurring with the PSF.

The Tharsis region, shown with a black rectangle in figure 5.1, is presented in more detail in figure 5.6. In the reconstruction in the right panel of figure 5.6, the line of the Tharsis Montes is seen to separate regions of higher and lower WEH abundances, where the north-west side of the line is enhanced in hydrogen with abundances $>$7 wt% WEH on the north-west flanks of the Tharsis Montes. This hydrogen concentration is greater than that in the raw data (figure 5.6 left panel) and than in Elphic et al. (2005a). The area with the highest WEH concentration is north-west of Olympus Mons (at 18.65° N, 226.2° E) and coincides with the Lycus Sulci region where the WEH content reaches 11 wt. %.

In addition to these hydrogen enhancements, Forget et al. (2006) suggest that ice should condense west of Elysium Mons (25.02° N, 147.21° E) during periods of high orbital obliquity. In figure 5.1 we also find a decreased epithermal neutron count rate, implying enhanced WEH, at this location. Forget et al. (2006) also note that if the source for the atmospheric water during high obliquity were not only the north polar cap but included a hypothetical south polar water cap, then ice would be deposited in the Eastern Hellas basin (40°S, 110°E). However, we see no evidence for this in figure 5.1.

5.2.3 Implications of the observed hydration

That we see a peak in the count rate in the reconstruction so closely coincident with the position of Olympus Mons in MOLA topography (figure 5.6) strengthens our faith in the other features shown in the reconstruction. The region of enhanced hydrogen content around Ascraeus Mons (12° N, 107° W) is larger than the glacial deposit identified by Parsons & Head (2005). This may be an interesting result, although establishing its significance will require determining the precise spatial resolution of the reconstruction at this location.

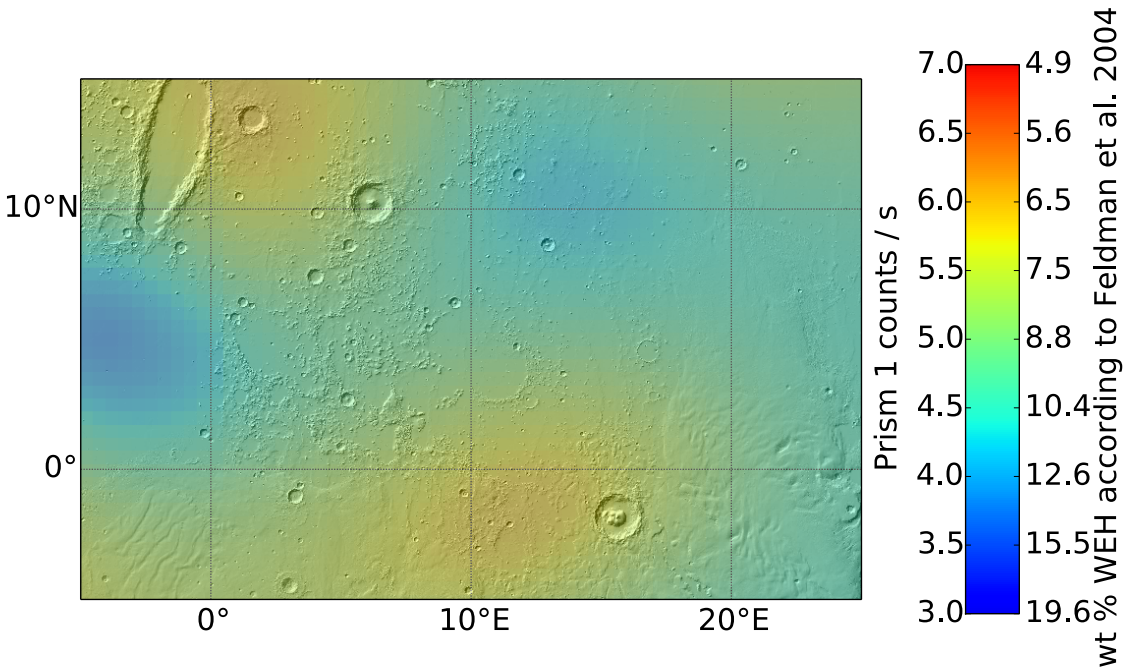

Figure 5.7: Locally adaptive pixon reconstruction of the MONS epithermal data around Meridiani Planum. Underlayed is a MOLA shaded relief map.

5.2.4 Meridiani Planum

Meridiani Planum is a plain centred at 0.2° N, 2.5° W and near the dichotomy boundary (the line separating the Martian northern lowlands and southern uplands). The plains became an area of intense interest after they were identified as being one of the few locations containing, and the only large deposit of, crystalline, grey hematite on Mars (Christensen et al., 2001). The favoured explanation for the origin of this mineral is chemical precipitation from aqueous Fe-rich fluids. This process requires the presence of liquid water, on the surface of Mars, for substantial periods of time, early in its history (Christensen et al., 2001). For this reason the region was selected as the landing site for NASA's Mars Exploration Rover, Opportunity.

The grey hematite deposits are centred on 2° S, 4° W. Comparison with figure 5.7 shows this area is not enriched in H in the MONS reconstruction. However, an area of the Planum ∼600 km to the north, at 4° N, 5° W, does show hydrogen enrichment. This region was shown, using the OMEGA visible-near infrared hyperspectral imager, to contain hydrated minerals including kieserite, gypsum, and polyhydrated sulfates (Gendrin et al., 2005). Why the hematite deposits appear less hydrogen rich than hydrated mineral

deposits to their north is not clear.

5.3 Desiccation due to Katabatic winds in Planum Boreum

The north pole is an interesting target for image reconstruction due to the presence of a perennial polar cap, made up mostly of water ice. As this feature is thought to be mostly water ice, it should give rise to a sharp edge in the spatial distribution of the neutron flux. Immediately surrounding and encircling the cap are a set of sand dunes, or ergs, including Olympia Undae, the largest continuous dune field on Mars (Tanaka et al., 2005), which stretches between 78° to 83° N and 155° to 230° E. Despite detailed morphological study (Lancaster & Greeley, 1990), the internal structure of these dunes is not precisely known. They possess a low thermal inertia (Paige et al., 1994), which is consistent with their being made up of cemented dust and sand (Greeley, 1986), uncemented water ice and silicate dust (Saunders & Blewett, 1987), or sand-sized andesite and basalt grains (Edgett et al., 2003).

The geological features at the north pole are illustrated in figure 5.8 and can be compared with the MONS data and pixon reconstruction in figure 5.9. It is clear in the unreconstructed MONS data that the polar cap is rich in hydrogen and the encircling dunes are poorer in hydrogen. Although the dunes are large, with Olympia Undae covering $\sim$ 300$\times$800 km, they are smaller than the PSF of the MONS instrument. However, with the improved resolution available in the reconstruction, this is not the case and the correspondence between geological features and water content is made much clearer. In the reconstruction the polar cap is revealed to contain 100 wt. % WEH over its extent and the dunes, drier than the surrounding plains, contain as little as 10 wt. % WEH. The variation in WEH abundance around the polar cap mirrors the dune coverage, with large areas dense with dunes, such as Olympia Undae and the region between 30° and 75° W, having a lower WEH abundance than regions with less dune coverage, such as 30° to 60° E, containing greater WEH concentrations.

A near pure water ice composition for the polar cap is in agreement with measurements from MARSIS (Picardi et al., 2005), sensitive to kilometre depths, and the Shallow Radar (SHARAD) on Mars Reconnaissance Orbiter, which is sensitive to vertical struc-

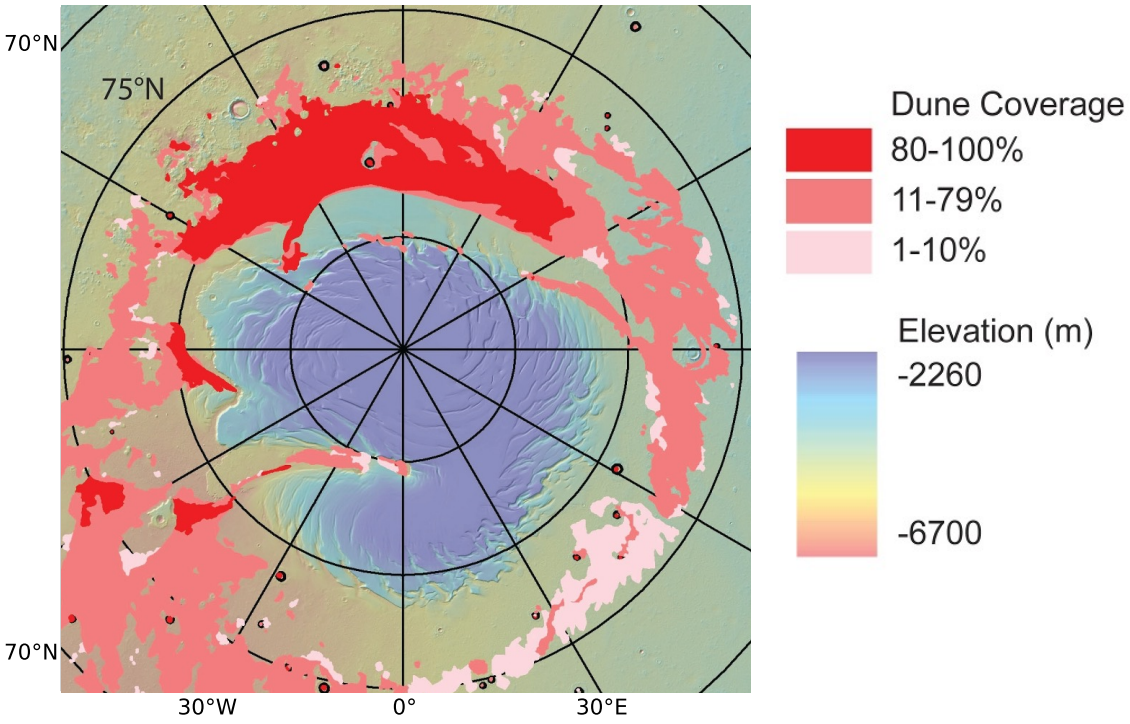

Figure 5.8: Stereographic projection of a topographic map based on MOLA data at the north pole, over which is plotted the fraction dune coverage, the polar cap is seen. Image adapted from Hayward et al. (2010).

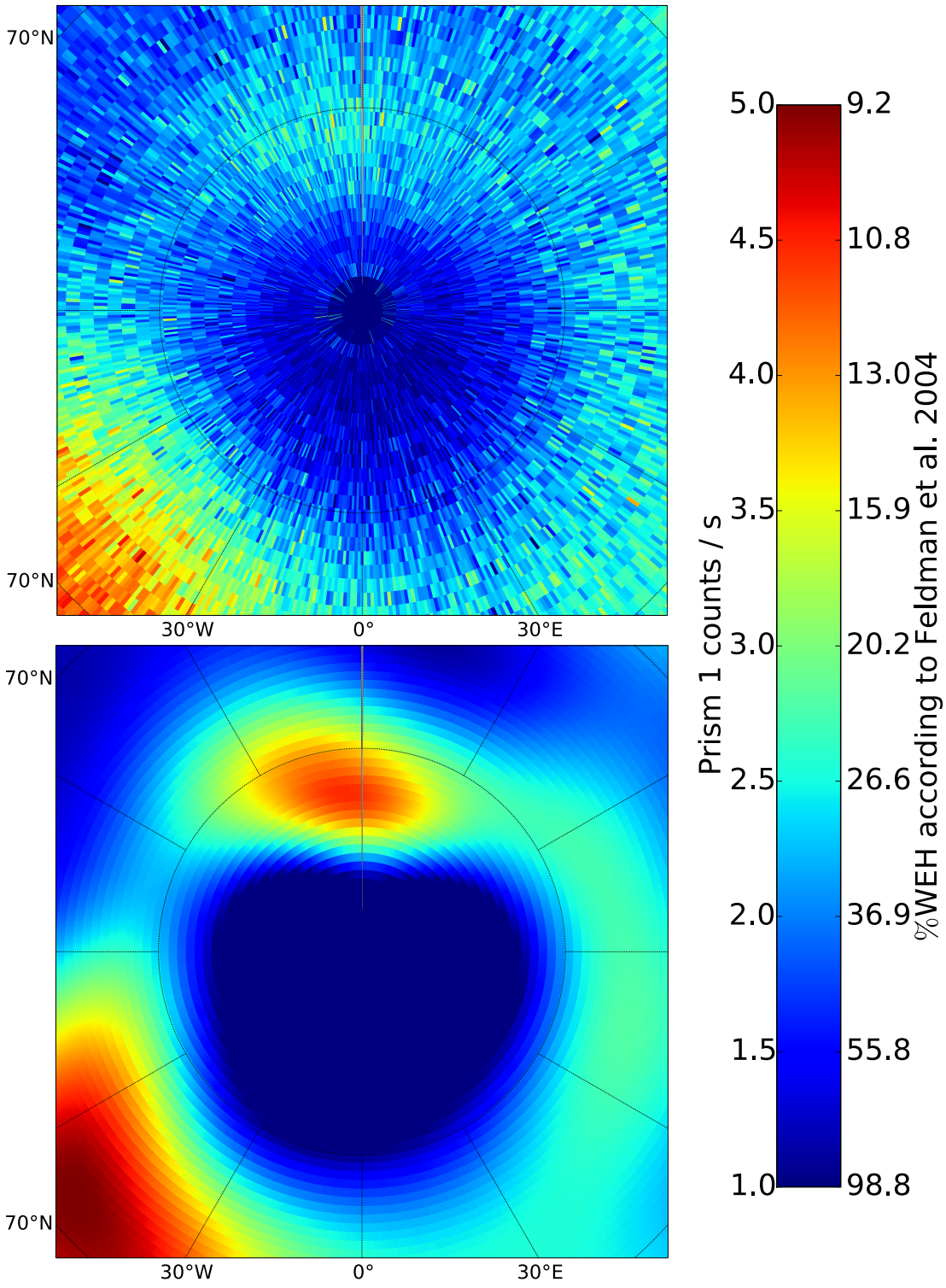

Figure 5.9: Stereographic map showing MONS data (*top*) and locally adaptive pixon reconstruction (*bottom*) covering the North pole and the northern dunes, including Olympia Undae.

tures on scales of at least 10 m (Phillips et al., 2008).

The relative desiccation of the ergs is the result of katabatic winds from the polar cap. These winds are driven by cold, high density air above the polar cap that warms adiabatically during their descent off the polar cap. Given the WEH abundances in the pixon reconstruction in figure 5.9, the composition of the dunes may be either a hydrated mineral such as gypsum or soil with pore-volume water ice emplaced by vapor diffusion overlying a desiccated layer.

Martian General Circulation Models (GCMs) depend on several important parameters associated with atmospheric and surface properties. One key property is the thermal inertia, which depends on the presence of water ice near the pole. Replicating the Viking and later missions atmospheric pressure histories requires taking into account near-surface water ice content and spatial distribution at high latitudes. In particular ice content is directly related to thermal conductivity and thermal inertia, and spatial variations of these govern the input and release of energy (and water vapour) seasonally. Therefore, the pixon reconstruction in the polar regions can be used to outline deviations from a uniform ice distribution poleward of 80° N which will influence local circulation and precipitation.

5.4 Recurring slope lineae

Recurring slope lineae are narrow, dark markings that appear and grow on Martian steep slopes during warm seasons and fade in cold seasons. They have been found mainly in the southern midlatitudes and were first identified using HiRISE images (McEwen et al., 2011). Contemporary liquid water activity on the surface of Mars has recently been confirmed at sites showing recurring slope lineae, using infra-red spectral data from the Compact Reconnaissance Imaging Spectrometer for Mars onboard the Mars Reconnaissance Orbiter (Ojha et al., 2015). The confirmed presence of liquid water on the surface of Mars has implications for astrobiology and, consequently, landing site selection. However, the origin of this fluid remains unknown.

The origin and recharge mechanism of these liquid brines detected in recurring slope lineae (RSL) has important implications for the current Martian water cycle and budget. However the exact mechanism that recharges the RSL, which may vary across the surface,

remains poorly understood. At least three possible scenarios exist:

1. Shallow briny aquifers are one possible source for the water in RSL. These must be within a few tens of centimetres of the surface as thermal inertia prevents seasonal melting at greater depths. Significant freezing point depression would be required to maintain the brine as a liquid so close to the surface. However perchlorate salts, which are observed in-situ (Hecht et al., 2009; Glavin et al., 2013) and at the locations of some RSL (Ojha et al., 2015), are able to suppress the freezing point of brines by up to 70 K (Chevrier et al., 2009). Ice and brines may be brought to the surface, from a deeply buried aquifer, by brine convection (Travis et al., 2013).

2. Shallow ice, melting during warm seasons, provides a good temporal match to RSL observations (McEwen et al., 2014) and shallow ice beneath flat terrain has been detected from new impact craters down to 39°N (Dundas et al., 2014). Such ice could provide a water source for the mid-latitude RSL by seasonal melting of relict frozen brines, deposited during periods of high orbital obliquity. The equator-facing slopes should quickly become depleted, and the continued presence of shallow frozen brines in equatorial regions is difficult to explain (McEwen et al., 2014). Annual recharge of subsurface water ice is not believed to be able to provide sufficient water to account for the RSL (Stillman et al., 2014; Grimm et al., 2014). However, deeper brines may be able to find their way to the surface in steep gullies and canyons especially deep in Valles Marineris (McEwen et al., 2014).

3. Finally, absorption of atmospheric water vapour by deliquescent salts could recharge RSL. However, questions remain over whether enough precipitable water vapour is present in the atmosphere to annually recharge the RSL (McEwen et al., 2014). Alternatively the water vapour may be supplied, not by the atmosphere, but from the continued sublimation of buried near-stable relict ice (Mellon et al., 1997).

To distinguish between these hypotheses, we use the results of a locally adaptive pixon reconstruction of the MONS epithermal neutron data to obtain knowledge of the near-surface water abundance.

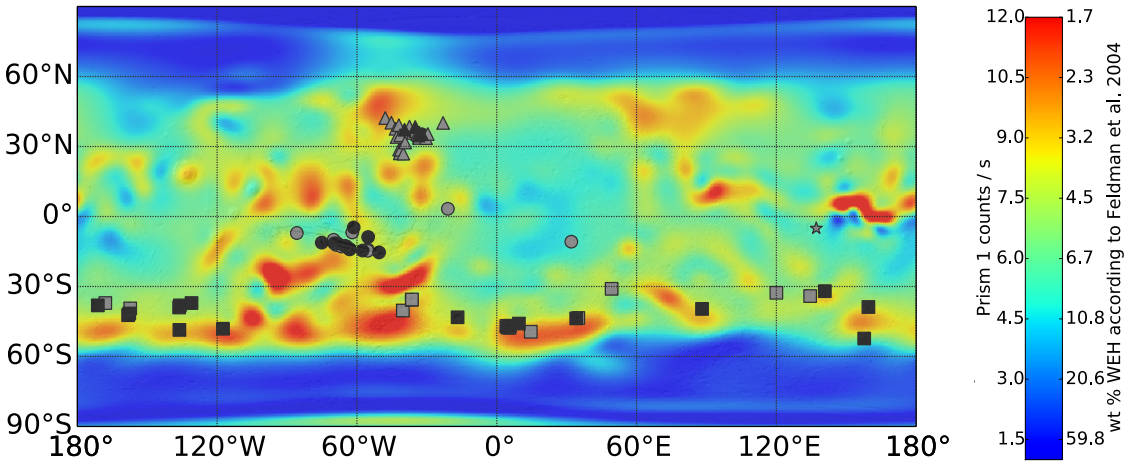

Figure 5.10: A global pixon reconstruction of the MONS prism-1 data. The colour bar also shows the conversion to wt. % water equivalent hydrogen from Feldman et al. (2004). The locations of both confirmed (black) and candidate (grey) RSL are also shown as identified in McEwen et al. (2011) (squares), Stillman et al. (2014) (circles) Stillman et al. (2016) (triangles) and Dundas & McEwen (2015) (star).

5.4.1 Correlation of RSL position with H abundance

Comparison of the high resolution hydrogen map with the locations of RSL shows no positive correlation (figure 5.10). To, quantitatively, test this comparison and find whether RSL occur in special regions in the WEH map we will examine the probability distribution of the count rate. Figure 5.11 shows the probability density function of the reconstructed count rate in the southern mid-latitudes (SML), along with the probability distribution of the count rate restricted to sites within the SML containing RSL. We compared these two distributions using a two-sample Kolmogorov-Smirnov test (Press et al., 1992), which measures the maximum difference between the cumulative distribution functions of two samples. This tests the null hypothesis that the data sets are drawn from the same distribution. Large values of the test statistic imply that the two samples are not drawn from the same underlying distribution. To carry out this calculation the data were repixelated more coarsely so that independent pixels are effectively uncorrelated. We find a Kolmogorov-Smirnov test statistic of 0.2 and corresponding p-value of 0.2 (i.e. test statistics at least this extreme are expected 20% of the time). This is not sufficient to imply that the count-rates (and consequently water abundances) where RSL are present are different from those locations where RSL are absent. In turn this result implies that the occurrence of RSL is

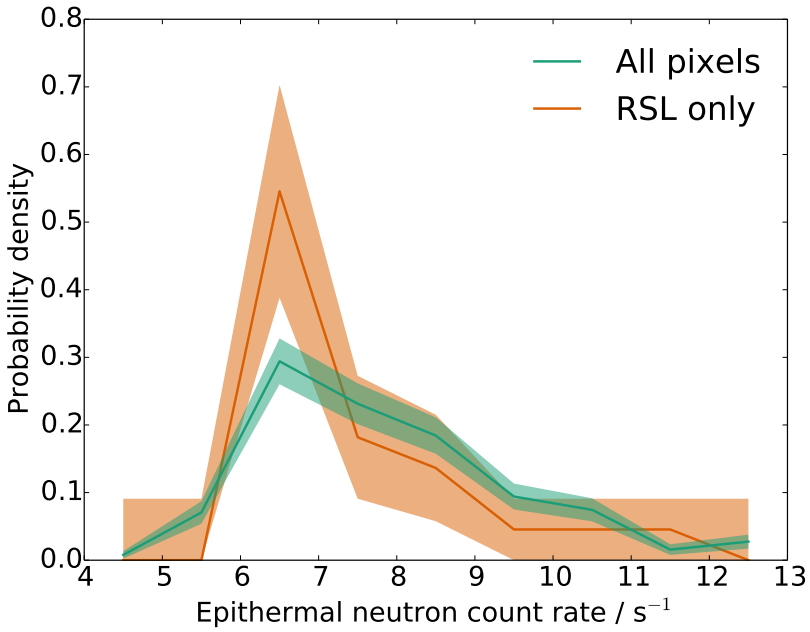

Figure 5.11: The probability density function of the count rate between 53°S and 30°S and just those pixels containing RSL. The translucent regions indicate 1σ errors.

independent of the WEH of the top few 10s of cm of soil on large scales. It should be noted that this analysis assumes that the distribution of discovered RSL reflects their true distribution across the surface in an unbiased way. If this assumption is incorrect and there are strong selection effects that correlate with WEH abundance then the results may change, for example if the low albedo regions in which RSL are found also have anomalous hydrogen abundances.

That WEH abundance in the soil is not correlated with RSL activity can be explained either by reference to the resolution of the neutron data, even at 290 km the resolution is too coarse and the regions containing RSL are blurred into those without, or we must conclude that there are no sizeable stores of near-subsurface water close to the RSL thus their formation and renewal is not driven by subsurface aquifers or water ice. Water buried beneath the level accessible to the MONS data (i.e. deeper than about 1 m) could not act as a source for the RSL because thermal inertia prevents ice buried beneath 20 cm from ever being melted (Stillman et al., 2014).

If the RSL water were sourced from a subsurface reservoir too small to be seen in MONS data then we may expect that improving the resolution would go some way to-

wards alleviating this discrepancy. The upper panels in Figure 4.11 show the MONS data around Valles Marineris before and after reconstruction and we find no evidence of a systematic increase in water abundance at the RSL with reconstruction. It is interesting to note that so far all of the confirmed RSL lie in the dry half of Valles Marineris east of Ius Chasma. Also, fluctuations apparent only in the reconstructed count rate maps match well with topographic features, e.g. eastern Valles Marineris and Eos Chasma appear dry compared to their surroundings. If the water source for the RSL is not subsurface then it must be either atmospheric or the sublimation of deeply buried ice deposits. But, both of these mechanisms require that the deliquescence of chlorate and perchlorate salts (the presence of which has been confirmed by CRISM and in-situ measurements (Hecht et al., 2009; Glavin et al., 2013; Ojha et al., 2015)) produces sufficient quantities of liquid brines to darken the soil and produce the RSL.

5.4.2 Limit of detection of increased hydration

To place a limit on the smallest source region that could be detected using the new pixon method we created a series of mock data sets containing circular regions of 100 wt% WEH with various diameters, in regions where RSL have been observed, i.e. the SML, Valles Marineris and the northern mid-latitudes. We then performed reconstructions of these mock data sets and examined the reconstructions to see how large the 100 wt% source region must be before a statistically significant depression in the count rate is seen. The variation of count rate with source size is shown in figure 5.12, where the error bars show the scatter in a set of ten mock data sets. It can be seen that sources smaller than approximately 60 km in radius do not lead to statistical significant decreases in count rate so are not reliably resolved using this method. Therefore, we cannot rule out near-subsurface water sources smaller than 120 km in diameter, though this is two orders of magnitude larger than the scales over which RSL occur.

5.5 Conclusions

We have investigated several regions of interest across the Martian surface that have previously been hypothesised to contain water or hydrated minerals. A consequence of blur-

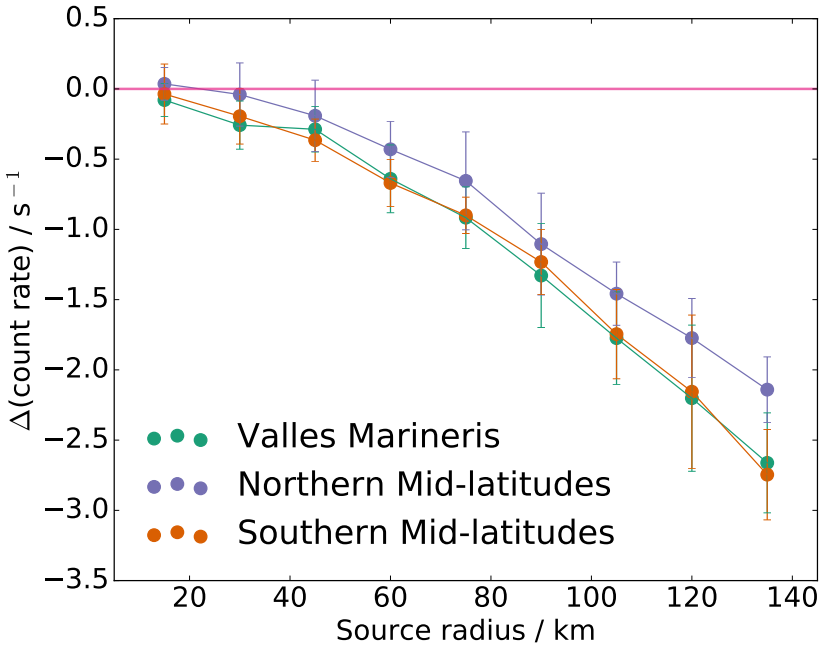

Figure 5.12: The difference in count rate between reconstructions with 100 wt% WEH sources and those without as a function of source radius. 1σ error bars are calculated from the standard deviation of 10 reconstructions with different noise realizations. The colours correspond to the regions where the sources were inserted. The horizontal line is at count rate difference equal to zero.

ring by the detector footprint is that earlier analysis of the neutron count rates had led to underestimation of the dynamic range of hydrogen abundance in regions with local variation.

At low and equatorial latitudes we found evidence, in the reconstruction of the MONS data, for buried water ice in the Medusae Fossae Formation and on the western slopes of the Tharsis montes and Elysium Mons. This supports the hypothesis that Mars in the recent geological past rotated on an axis highly inclined to its current one. On the other hand, the region containing plate-like features and previously to be a buried water ice sea was found to be exceptionally dry. Although this region seems not to contain hydrogen within the top few tens of cm of the surface, its unusually high epithermal neutron flux may indicate a unique composition. Further work using Monte Carlo particle transport codes is required to learn more about the composition of this feature.

Recurring slope lineae are transient features recently discovered on Martian steep slopes during warm seasons. Their narrow, dark shapes resemble water courses – and the presence of liquid water has recently been spectroscopically confirmed near some RSL sites. Yet, if this is what they are, then their source remains unknown. We find that the hydrogen abundance at RSL sites is statistically no different from anywhere else at similar latitudes. Indeed, some RSL are found in very dry locations. This implies that RSL are not supplied by large, near-surface briny aquifers, as had been a leading theory. Preferred hypotheses remain for RSL to be fed by water vapour, either from sublimating deeply buried water ice or from the atmosphere.

CHAPTER 6

Time dependent variation in the neutron data and the Martian CO_2 cycle

The Martian polar regions are key to understanding the current climate and energy balance, with more than 25% of the Martian atmosphere participating in the annual polar CO_2 sublimation and condensation cycle (Hess et al., 1979, 1980; James et al., 1992). Presently, both poles contain permanent water ice caps several km in depth (Carr, 1996), with that at the south pole overlain by a permanent cover of CO_2 frost around 1 m thick (Kieffer, 1979). In addition to these perennial water ice caps gamma ray and neutron data have shown that substantial subsurface water ice deposits are present down to around $\pm$ 60° (Feldman et al., 2002; Mitrofanov et al., 2002). On top of the caps and down to $\pm$ 60° seasonal CO_2 frost condenses and sublimates from the atmosphere between autumn and spring. From atmospheric pressure measurements it is estimated that 3.5×10^{15} kg of CO_2 condense in the north, and 8.1×10^{15} kg of CO_2 in the south to form the seasonal polar caps (Kieffer et al., 1992).

Much of our knowledge of the Martian atmosphere is encapsulated in the General Circulation Models (GCMs): numerical simulations based on those developed for weather and climate forecasting on Earth (Leovy & Mintz, 1969; Forget et al., 1999). These models are constrained using the in-situ measurements of the Martian atmosphere made by the various landers and rovers from the Viking missions (Forget et al., 1999) to the present MSL Curiosity rover (Harri et al., 2014). However these measurements lack either the spatial and temporal resolution or spatial coverage to fully characterize the Martian atmosphere.

The predictions of the GCMs have been compared with remotely sensed global data including that from the Thermal Emission Spectrometer (TES), the Mars Climate Sounder

(Navarro et al., 2014) and the Mars Reconnaissance Orbiter's Compact Reconnaissance Imaging and Spectral Mapping (CRISM) instrument (Clancy et al., 2012). The GCMs are found to be in agreement with most available observations and have been used to predict previously unobserved phenomena, including meridional circulations (Haberle et al., 1993) and planetary waves (Barnes et al., 1993).

Mars Odyssey gamma ray and neutron data have also been used to infer the abundance of CO_2 at the poles (Feldman et al., 2003; Kelly et al., 2006; Litvak et al., 2006, 2007; Prettyman et al., 2009). CO_2 is a much less effective moderator of epithermal neutrons than is hydrogen. Thus, as CO_2 is deposited during autumn and winter the epithermal neutron flux observed above the Martian poles increase. Modelling this process using particle transport codes enables a relation between the seasonal change in count rate and amount of CO_2 deposited to be found. Seasonally deposited CO_2 also attenuates those gamma rays emitted from the Martian surface. Observation of the seasonal attenuation of the hydrogen-neutron capture gamma ray line, which provides a strong signal at the Martian poles due to their high hydrogen abundance, can also be used infer the presence of CO_2 frost at the poles. Previously, the utility of these data sets in constraining the distribution of water and CO_2, on scales similar to the GCM resolution (typically $\sim$ 300 km), has been limited by their poor spatial resolution due to the large footprint of the detectors (the MONS PSF has a full width at half maximum of 520 km and the others are comparable). This spatial smoothing will lead to errors in estimation of composition where the count rate varies on scales smaller than the PSF, such as at the boundary of the CO_2 cap. Prettyman et al. (2009) used Jansson's method to remove the blurring due to the PSF but this method is known to perform poorly in the presence of noise (refer back to chapter 2 for details).

In addition to gamma ray and neutron data, the mass of CO_2 at the poles can be inferred from topography, using the Mars Orbiter Laser Altimeter (MOLA) (Smith et al., 2001), or using energy balance calculations based on Mars Global Surveyor's Thermal Emission Spectrometer (TES) observations of the solar and infrared radiation fields (Kieffer & Titus, 2001; Titus et al., 2003). However, of these three methods gamma ray and neutron spectroscopy is believed to be the most accurate and direct measurement (Haberle et al., 2004), as those based on MOLA data require the density of the cap to be known,

which it is not (Smith et al., 2001), and TES based inferences have neglected the role of subsurface heat conduction (Titus et al., 2003). Also, during the springtime recession of the cap, visible imaging (Mount & Titus, 2015) and visible and near-infrared spectroscopy (Appéré et al., 2011; Brown et al., 2012) can be used to track the retreat of the cap, though these methods are sensitive only to the abundance in the top few μm (during deposition the cap is not illuminated by the Sun so cannot be imaged).

In this chapter, we will apply the pixon image reconstruction technique to MONS epithermal data to infer the changing CO_2 frost abundance at the Martian poles and look for signs of non-polar temporal variation. This is the first application of the technique, on a global scale, to data split seasonally. We will look for non-polar variation in the neutron count rate, which could be indicative of seasonal water transport across Mars, along with the polar data, which will be used to infer the form of the seasonal deposition and sublimation of CO_2 frost.

6.1 Time varying data

To find temporal changes in our reconstruction we must split the data temporally. However, the resolution increase achieved by the pixon methods is dependent on the amount of data that is used in the reconstruction. We, therefore, would like to use as few bins as possible, while still being able to achieve reasonable temporal resolution. Additionally, as resolution is dependent on the amount of data, having each bin contain the same number of observations will help to reduce the difference in resolution between reconstructions made at different times. Empirically we have settled on splitting the MONS data into 12 bins in Martian solar longitude containing equal numbers of observations with each of these data sets pixelated in 1° equal angle maps when a high spatial resolution is desired; for example when examining the shape of features in the south polar cap. When comparing with the GCM, which is defined in a 5° equal angle grid, 60 solar longitude bins are used, giving finer temporal resolution. These binnings have the effect of averaging together separate Martian years so cannot be used to find secular evolution; the binning is illustrated in figure 6.1. The count rates vary with seasonal CO_2 frost thickness, increasing in autumn and winter and decreasing in spring in response to changes in insolation. In

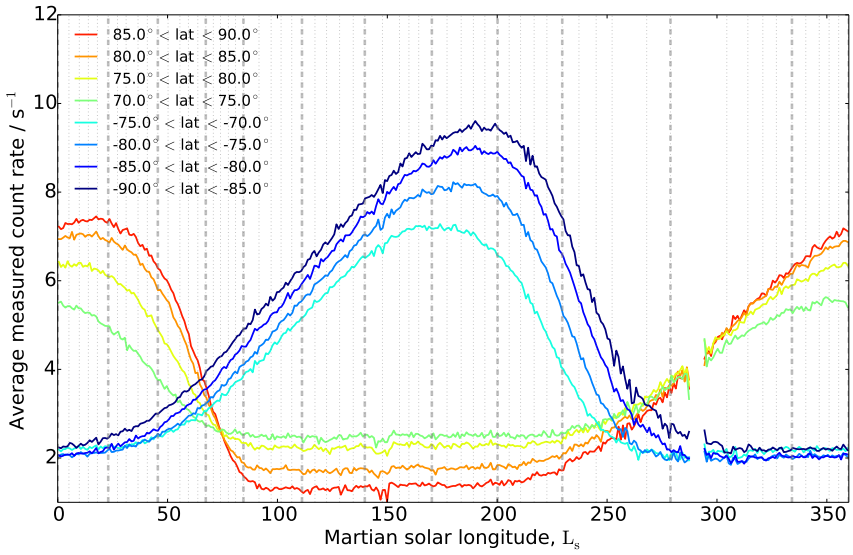

Figure 6.1: The average measured epithermal neutron count rate, split into a series of near-polar latitude bands, for given values of Martian solar longitude, L_S. The dashed grey lines show the boundaries of the solar longitude bins into which the data have been split for the different whole surface reconstructions in figures 6.3 and 6.5. The dotted grey lines show the temporal binning used for the reconstructions that went into making figures 6.6 and 6.7.

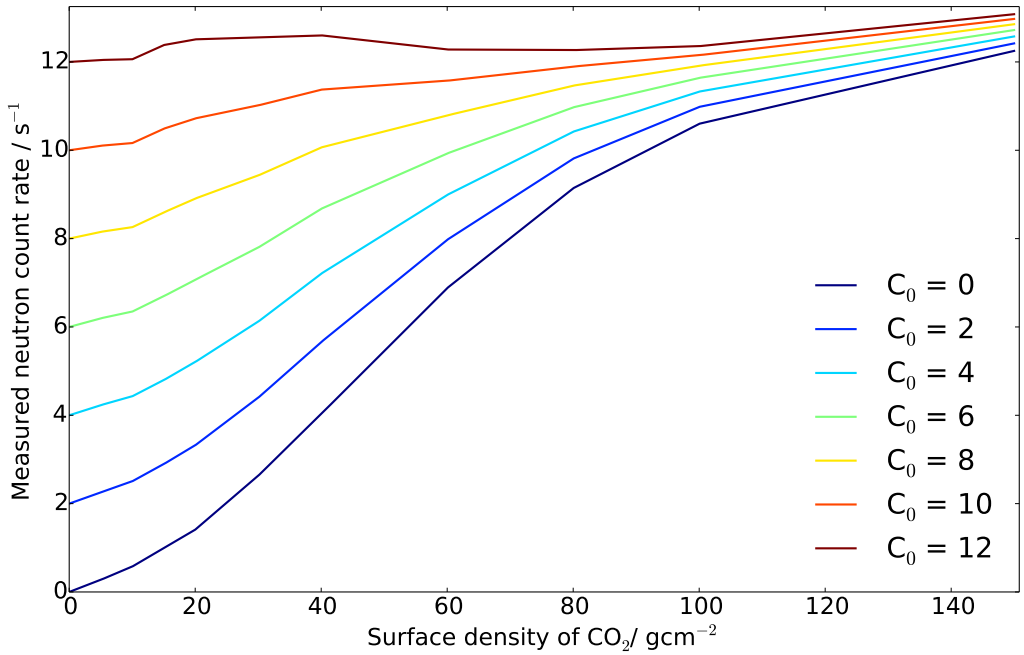

Figure 6.2: Fits showing how the measured epithermal neutron count rate, C, varies when a given amount of CO_2, A, is deposited on the surface, for a range of frost-free, summertime count rates, C_0.

winter, the low temperatures at the poles allow for the condensation of CO_2 from the atmosphere. As CO_2 is a poor moderator of epithermal neutrons this increase in CO_2 cover leads to an increase in epithermal neutron flux, compared with that when the surface consists of ice or ice-rich soil. It should be noted that in figure 6.1 the count rate varies by up to 5 counts s^{-1} within a single bin at the south pole during spring. This will increase the errors in each pixel which are, as in chapter 4, estimated empirically by calculating the scatter between repeat observations in the observed count rate within the pixel. Consequently, the reconstructions in the bins that contain these large count rate changes, will be more uncertain.

6.2 Connecting neutron count rate with CO_2 abundance

In this chapter we will use the results of Prettyman et al. (2009) when converting from a measured epithermal neutron count rate to a surface density of CO_2. They modelled the Martian poles as a semi-infinite surface of constant WEH over which CO_2 frost is deposited seasonally. Using a Monte Carlo transport code they devised an interpolation

Table 6.1: Parameters used in the epithermal neutron count rate to CO_2 surface density conversion.

	South Pole			North Pole		
A(g cm^{-2})	a_0	a_1	a_2	a_0	a_1	a_2
0	0	1	0	0	1	0
5	0.173	1.03	-0.00556	0.274	0.993	-0.00101
10	0.369	1.03	-0.00893	0.574	0.967	-0.000828
15	0.640	1.03	-0.0116	0.991	0.952	-0.000229
20	0.926	1.07	-0.0176	1.41	0.961	-0.00299
30	1.86	1.07	-0.0240	2.63	0.895	-0.00565
40	3.07	1.04	-0.0296	4.03	0.834	-0.00997
60	6.10	0.800	-0.0267	6.88	0.566	-0.00966
80	8.76	0.517	-0.0190	9.14	0.350	-0.00745
100	10.7	0.297	-0.0107	10.6	0.200	-0.00447
200	15.1	-0.0574	0.00450	13.9	-0.0239	0.00133

scheme relating the frost-free, summertime epithermal neutron count rate at a point on the surface, C_0, and surface density of CO_2 at that location, A in g cm^{-2}, to the measured count rate at a given time of year, C,

$$C(A,C_0) = a_0(A) + a_1(A)C_0 + a_2(A)C_0^2, \tag{6.2.1}$$

where the parameters a_0, a_1 and a_2 are given in table (6.1). The value of the parameters in equation (6.2.1) were found to depend on the atmospheric column density, which is strongly dependent on topography, so are different for the north and south poles. This relation can be understood as the result of CO_2's low cross-section for epithermal neutron capture, compared with hydrogen, leading to less moderation with increasing CO_2 abundance.

For our purposes, equation (6.2.1) must be inverted to give A for measured C and C_0. The equation is shown for a range of C_0 in figure 6.2. It can be seen that for summertime count rates below 12 counts s^{-1} a monotonic and unique mapping exists between surface density of CO_2, A, and count rate, C, so the function can be trivially numerically inverted. For frost-free count rates above 12 counts s^{-1} this is not true. However, if we restrict our attention to within 30° of the pole these values will not be encountered. Also evident in figure 6.2 is that the sensitivity to changes in CO_2 abundance decreases with increasing count rate, or equivalently with decreasing WEH abundance.

Prettyman et al. (2009) note the presence of several effects that act as systematic errors on this conversion. These include: vertical variation in soil WEH abundance, which they estimate, via a set of simulations with different WEH abundances, to lead to a <7 gcm^{-2} error; and the concentration of non-condensable, neutron absorbing gasses (N_2 and Ar) in the atmosphere, which, for CO_2 deposits similar in size to those at the poles are believed to be, could lead to an error of 20% in the estimate of CO_2 surface density.

6.3 Reconstructions

To suppress noise and improve the resolution of the data we have performed a locally adaptive pixon reconstruction on each of the 72 different solar longitude data sets defined in section 6.1. The north and south poles of four of the high spatial resolution reconstructions are shown in figure 6.3. The seasonal changes in count rate away from the poles are much less noticeable and are not illustrated here.

As in section 4.3.1 the improvement in spatial resolution achieved by the reconstruction process is estimated by using a two point angular correlation function for the data and reconstructions, shown in figure 6.4. It can be seen from the figure that despite each data set having the same number of observations there is still some scatter in the resolution achieved, which varies between 270 and 350 km. This represents a global average, however, and due to the large number of observations at the poles the pixon width, which is the scale over which the method allows the reconstruction to vary, drops to 100 km within 10° of the poles. Thus, the actual resolution at the poles may be somewhat less improved than the global value suggest.

The spatial resolution of the 60 low spatial resolution reconstructions are limited by the pixelation, so they can be taken to have a resolution of 5°.

6.4 Details of the CO_2 cap advance and retreat

The abundance of CO_2 can be inferred from MONS data using the procedure described in section 6.2. Applying this technique to our improved resolution reconstructions allows for a more effective comparison of the GCMs to observations than using the raw MONS

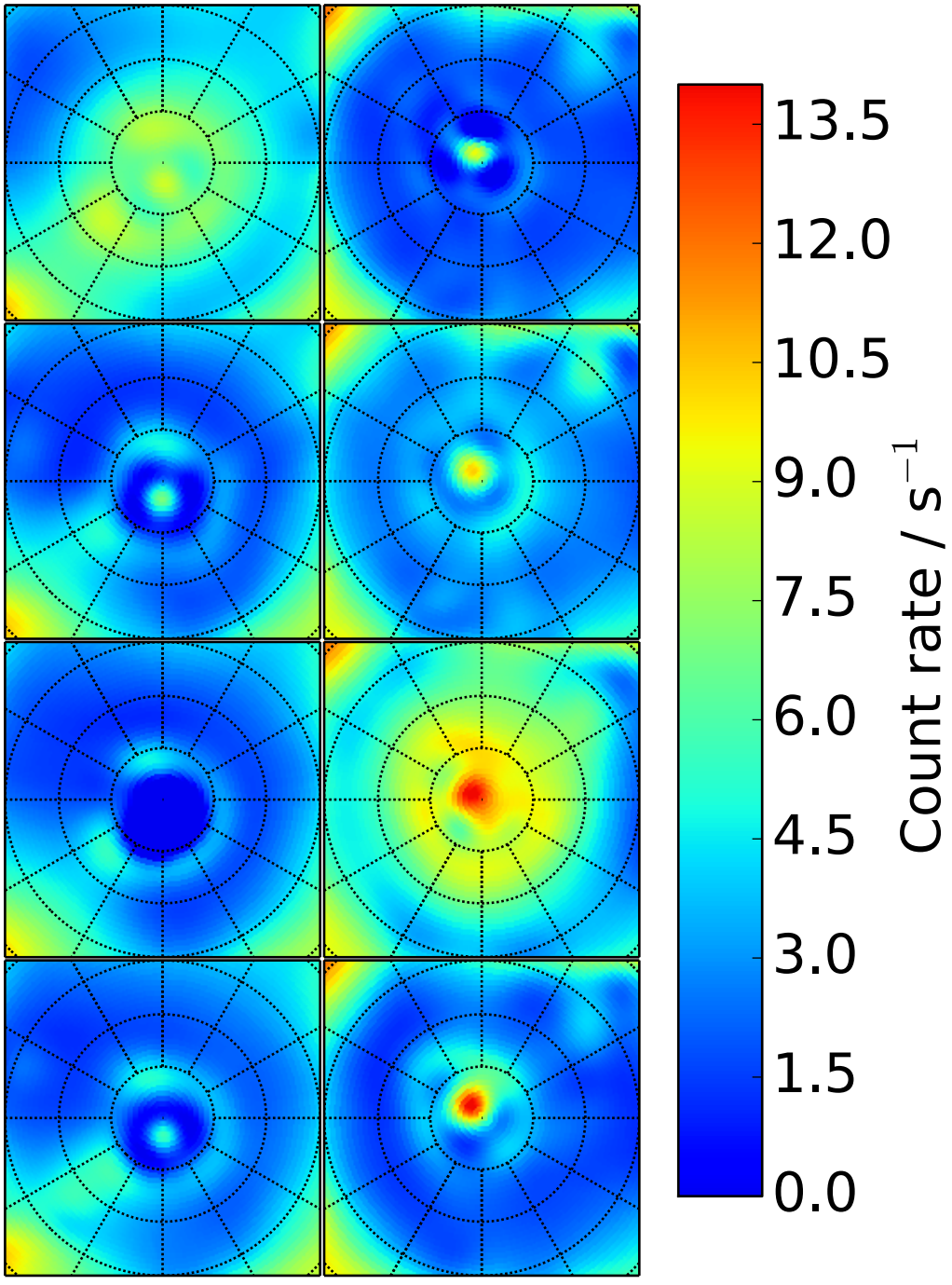

Figure 6.3: Polar stereographic projection of the locally adaptive pixon reconstructions of the MONS epithermal data at the north (*left*) and south (*right*) poles, going out to 30° from the poles. From top to bottom the solar longitude bins are centred on 347°, 76°, 185° and 254°. Grids indicate every 30°of longitude and every 10°of latitude. For the north pole 0°W longitude is at the bottom of the figure, and W longitude increases clockwise from that point, in the south 0°E longitude is at the top of the figure, and E longitude increases clockwise from that point. The subplots are 60°or 3550 km across.

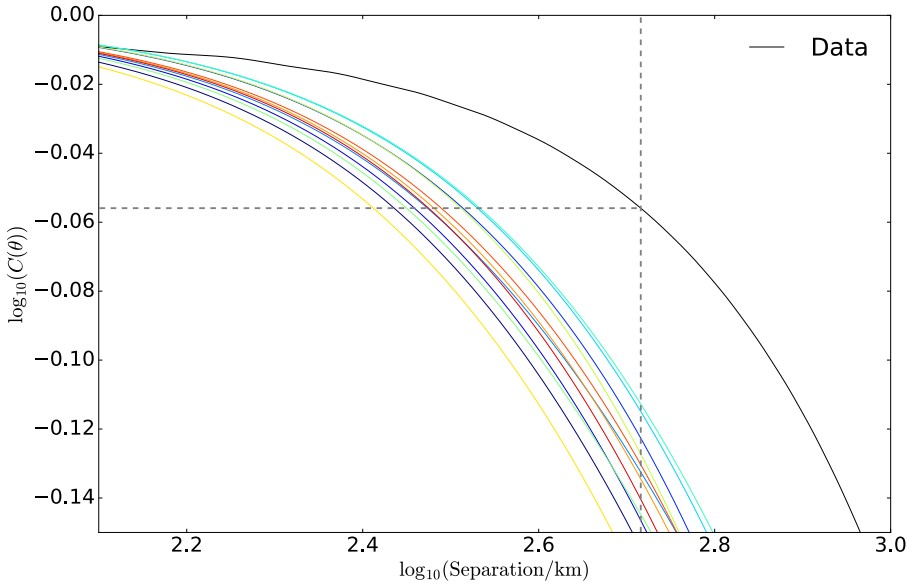

Figure 6.4: The two point angular correlation function of the MONS data and the set of 12 different solar longitude high spatial resolution reconstructions, normalized to unity at zero separation. The vertical dashed line shows the MONS PSF FWHM scale. The dashed vertical line shows the separation required for the power in the reconstructions to drop by the same proportion that the power in the data drops at a separation equivalent to the FWHM of the PSF.

data. When performing this conversion we have taken the summertime count rates, C_0, to be those in the reconstructions of the data between Martian solar longitudes of 140° and 170° in the north and 334° and 360° in the south. Consequently, these reconstructions are defined as containing no CO_2, as illustrated in the top right panel of figure 6.5, which shows the CO_2 surface density inferred from selected reconstructions. This approximation is obviously incorrect at the south pole, which is known to contain a permanent CO_2 cap (Kieffer, 1979), and will lead to an underestimate of the CO_2 abundances within the 5° diameter perennial south polar cap. However, as the permanent CO_2 cap has less than 1% of the area of the seasonal cap the effect of this error on the total abundance of CO_2 should be negligible.

In the high spatial resolution reconstructions at both the north and south poles the growth of the CO_2 cap is not centred on the geographic pole. At the south pole the seasonal cap growth starts and reaches its greatest thickness coincident with the perennial cap. At the north pole this point occurs at 85° N, 0° E, near the end of the Chasma Boreale. The offset at the south pole has been seen in other studies using MONS neutron data

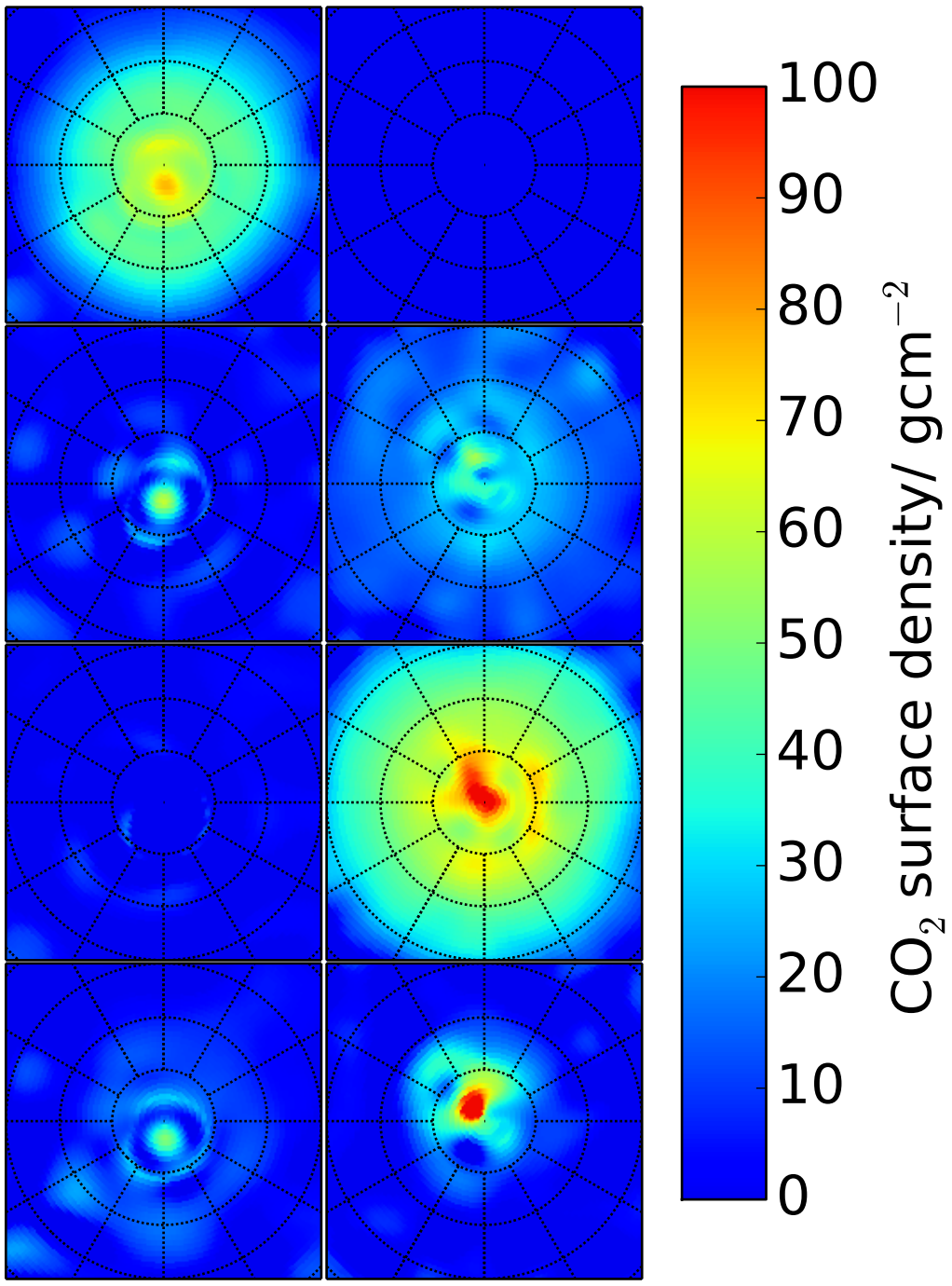

Figure 6.5: Polar stereographic projection of the inferred CO_2 surface density from locally adaptive pixon reconstructions of the MONS epithermal data at the north (*left*) and south (*right*) poles, going out to 30° from the poles. From top to bottom the solar longitude bins are centred on 347°, 76°, 155° and 254°. Grids indicate every 30°of longitude and every 10°of latitude. For the north pole 0°W longitude is at the bottom of the figure, and W longitude increases clockwise from that point, in the south 0°E longitude is at the top of the figure, and E longitude increases clockwise from that point. The subplots are 60°or 3550 km across.

(Prettyman et al., 2009), but that at the north pole has not. It should be noted that in our 60 low spatial resolution reconstructions the offset is seen at the south but not at the north pole, in agreement with earlier work (Prettyman et al., 2009). This result is consistent with the appearance of the offset in the north polar seasonal cap being the result of the increased resolution available in our new reconstructions. Modelling has shown that the offset at the south pole is due to the effects of topography on local weather (Colaprete et al., 2005); the nearby large basin, Hellas, results in a low pressure system, favouring the deposition of CO_2 as snow rather than frost. The higher albedo of CO_2 snow leads to slower sublimation and a permanent CO_2 cap offset from the south pole (Colaprete et al., 2005). However, there is no similar impact basin close to the north pole, so a similar offset in CO_2 deposition is not predicted.

The higher resolution of the reconstructions that were used to infer the CO_2 abundances shown in figure 6.5 allows structure in the CO_2 cap to be determined. Again, this detailed structure is not visible either in our 60 high temporal resolution reconstructions or in previous work (Prettyman et al., 2009). Imaging is not available during the deposition of the cap, but OMEGA visible / near-infrared imaging spectrometer observations show strong spatial variations in the CO_2 ice distribution during recession (Appéré et al., 2011). CRISM spectral imaging also shows inhomogeneity in the CO_2 distribution, along with the deposition of water over the retreating seasonal caps (Brown et al., 2010, 2012). If the inhomogeneities detected in these spectral data sets are present at depth, and not merely surfical, then they may be detected in the neutron data. However, it should be pointed out that although much of the structure seen in figure 6.5 persists between reconstructions of different time periods, it does not obviously correspond to structure seen in imaging data such as that shown in Appéré et al. (2011) and Brown et al. (2010). Also, the exact form of the spatial structure, but not the timing of the onset of deposition, seems to be dependent on the reconstruction technique used, as reconstructions based on the maximum entropy pixon reconstruction are different in detail to those shown in figure 6.5. However, the differences between the two techniques disappear when averaging over broad regions of the cap.

Figure 6.6 shows a comparison of a GCM and 60 high temporal resolution reconstructed polar CO_2 abundances, averaged over 5° latitude bands. The GCM used in the

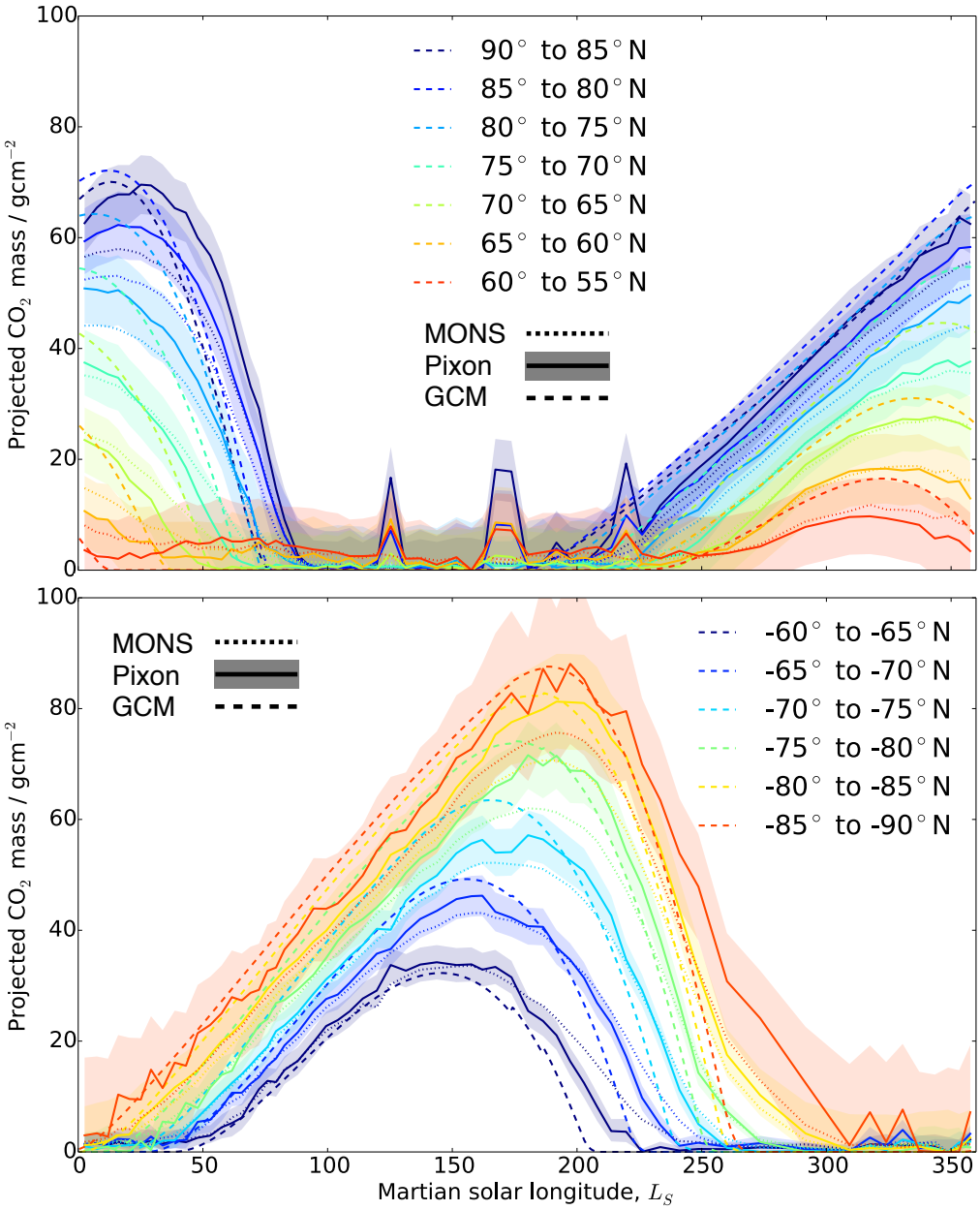

Figure 6.6: The variation of the abundance of CO_2, averaged in 5° latitude bands within 30° of the north pole (*top*) and south pole (*bottom*). The dotted and solid curves are values derived from the MONS data and a pixon reconstruction of that data, respectively, using the method of Prettyman et al. (2009) for conversion. The dashed lines are the results of a GCM and the different colours correspond to different latitude bins. The translucent region illustrates the 1σ errors on the CO_2 abundance inferred from the pixon reconstruction. To avoid confusion the errors on the MONS data curves are not shown, but are comparable in size to those for the reconstruction.

comparison is from Steele et al. (2014), which includes an assimilation of the TES data. That is, the GCM has been forced to agree with the TES measurements, which should produce results that better model the Martian water cycle than do unassimilated reconstructions. As the soil's water abundance determines its thermal inertia water abundance has a strong influence on the rate of CO_2 deposition and sublimation. In the figure we see broad agreement between the reconstructions and the GCM, including the amount of CO_2 deposited, the timing of deposition and sublimation and the relative amounts of CO_2 deposited in the north and south poles. However, there are differences in the details. At the south pole all points equatorward of 85°S̃ are found to contain less CO_2 than the GCM predicts, however the discrepancy is less than 1σ at all but one latitude, and the rate of sublimation in spring is lower. Also, poleward of 70°S̃ the peak in the amount of deposited CO_2 occurs later than predicted in the GCM, although the errors in the CO_2 abundance allow for agreement with the GCM the consistent presence of the offset suggests that it is not a random error. Other studies have not had the temporal resolution to determine whether a temporal offset between the GCMs and gamma ray/neutron data is present (Kelly et al., 2006; Prettyman et al., 2009). At the north pole the disagreement is more pronounced, with all latitudes south of 80° N containing at least 1σ less CO_2 than predicted. Above 80° N our data agree with previous studies. We obtain a peak CO_2 abundance in the 80° N $<$ lat $<$ 85° N band of 62 ± 6 g cm^{-2}, compared with 64 ± 5 g cm^{-2} in Prettyman et al. (2009), and 62 ± 5 g cm^{-2} based on gamma ray data (Kelly et al., 2006). The reason for the discrepancy at lower latitudes is not clear, however at these latitudes the summer and wintertime count rates are closer than at the poles, so errors in the estimation of the CO_2 free count rate, C_0, will lead to greater offsets in our estimates of CO_2 abundance.

Comparing the reconstruction and unreconstructed MONS data in figure 6.6 we see that the CO_2 abundance inferred from the reconstructions is larger than that inferred directly from the MONS data, except between 60° and 70°Ñ, and the recession happens more quickly. The reconstructions thus show better agreement with the GCM than does the MONS data. This is consistent with some of the discrepancy between the GCM and data being due to regions with different CO_2 abundances being smoothed together.

The total mass of CO_2 in the seasonal cap poleward of $\pm60°$ was determined, based

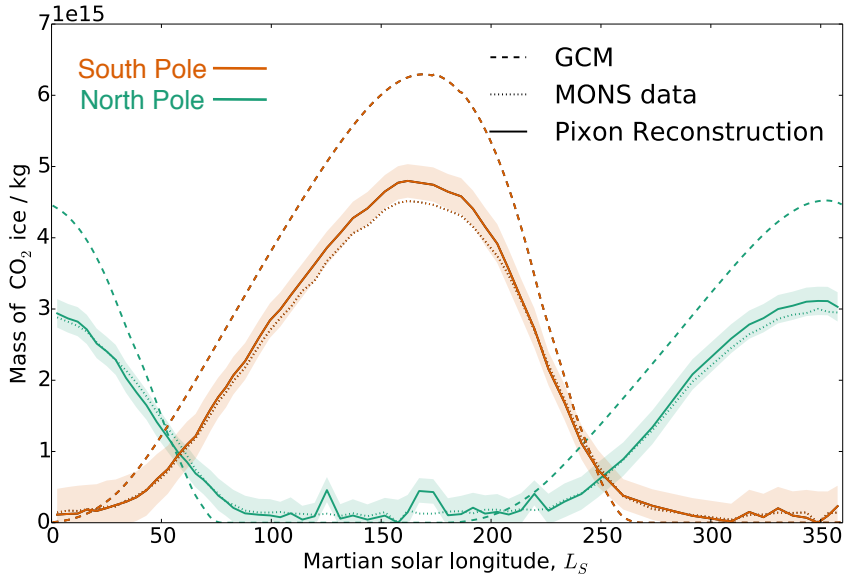

Figure 6.7: The total abundance of CO_2 within 30° of the north pole (green) and south pole (orange). The dotted and solid lines are values derived from the MONS data and a pixon reconstruction of that data, respectively, using the method of Prettyman et al. (2009) for conversion. The dashed lines are the results of a GCM. The translucent regions show the 1σ errors on the pixon curves.

on the MONS data and 60 high temporal resolution reconstructions, and is compared with the GCM in figure 6.7. Although the shape of all of the curves, including the timing of deposition, sublimation and peak CO_2 mass, is consistent, the total amount of CO_2 inferred from the neutron data is much lower than in the GCM. At the north pole the seasonal cap has nearly 50% more CO_2 in the GCM than is measured in the MONS data. Viking pressure measurements imply that 3.5×10^{15} kg of CO_2 condense in the north (Kieffer et al., 1992), somewhat less than predicted by the GCM, but still 20% more than measured in the neutron data. As measuring the total mass in the cap involves averaging over a large area, the amount of CO_2 measured in the MONS data and reconstruction are similar, agreeing almost everywhere to within 1σ. This result is unexpected as Prettyman et al. (2009), also using MONS epithermal neutron data, found agreement to within 10% between CO_2 mass based on MONS data and the NASA Ames GCM (Prettyman et al., 2004), which gives similar CO_2 mass predictions to the GCM used in figures 6.6 and 6.7. The MONS data used by Prettyman et al. (2009) and this study have followed independent data reduction pipelines and we do not have access to the data used by Prettyman et al. (2009). The disagreement between the results shown in figure 6.7 and previous studies may be due to differences in the data reduction and reflect an additional systematic errors in the CO_2 conversion relation.

6.5 Non-polar neutron count rate variation

In addition to the CO_2 cycle, GCMs predict seasonal water transport including the Martian atmosphere and near-surface soil (Steele et al., 2014). Also, evidence for seasonal hydration at RSL sites is seen in spectral data (Ojha et al., 2015) (see section 5.4 for more detail). Given our time varying reconstructions of the MONS epithermal neutron data we have the ability to look for such seasonal variations in hydrogen abundance, with the possibility of finding new evidence for a contemporary Martian water cycle.

To search for these seasonal changes in WEH we have calculated the standard deviation of the reconstructed count rate between all 12 different solar longitude data sets shown in figure 6.1, which is plotted in figure 6.8. In this figure, the largest scatter is seen at the poles, however there are some locations near the equator where large deviations in

the reconstructed count rate are seen, the structure in the figure reflects the structure in the reconstructed images and is set by the pixon distribution. This large, equatorial variation is shown, for a few locations, in figure 6.9. Given the 1σ uncertainties illustrated in this figure, the change in count rate is unlikely to be caused by random errors in the data or reconstruction process. However, there is an additional effect that can change the count rate locally: the changing pixon size. Although the data sets were each chosen to contain roughly equal numbers of observations, the distribution of these observations over the sphere is not constant. Thus in some data sets certain areas have fewer observations than in other data sets, which affects the pixon size. This, in turn, changes the count rate through the degree to which local peaks or troughs are smoothed with their surroundings. The orange curves in figure 6.9 correspond to a location at 32° S, 103° W. These count rates and pixon sizes are strongly correlated, which is consistent with the apparent temporal variation in the reconstructed feature being due to variation in pixon size in a small low-count rate region. The sharp peak in the violet count rate curve at $L_S \sim 80^\circ$ is coincident with a sharp decrease in the pixon size, suggesting this feature is surrounded by lower count rate regions. The green curves correspond to 1° S, 149° E, which is in Zephyria Planum in the Medusae Fossae formation. These count rate and pixon size curves are not obviously correlated, so the possibility remains that these count rate changes are due to a variation in WEH abundance, which would imply the seasonal transport of macroscopic amounts of hydrogen.

The correlation between pixon size and count rate makes automatically detecting real, temporal changes in the WEH abundance difficult. We have examined several other non-polar locations with high count rate scatter between the reconstructions, but so far have failed to find any (other than at the MFF) locations where the count rate variation is not correlated with changing pixon size and is co-located with a potentially interesting surface feature.

6.6 Conclusions

We have used a locally adaptive pixon reconstruction to examine the seasonal change in epithermal neutron flux. At equatorial and mid latitudes no convincing evidence for

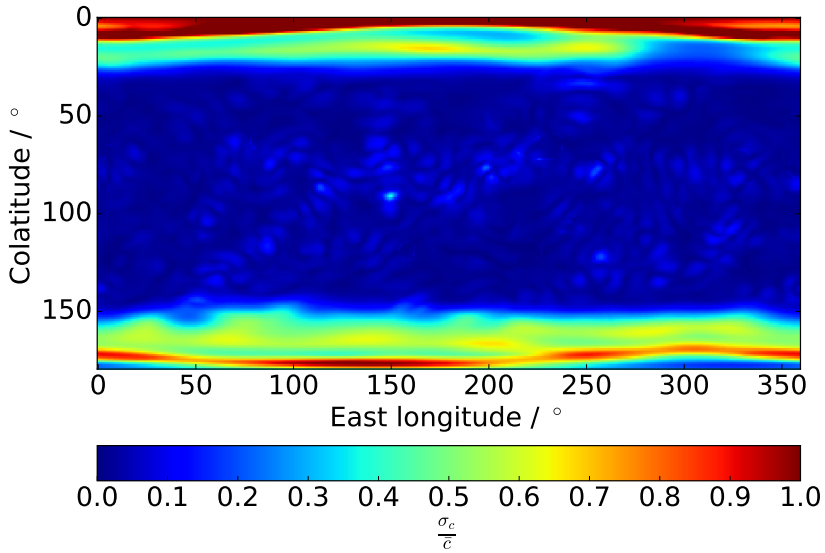

Figure 6.8: The standard deviation of the reconstructed count rate, across the 12 reconstructions, σ_c, normalized by the average value of the count rate, in each pixel, $\bar{c}$.

changing neutron flux driven by changing WEH content was found. However, determining which fluctuations in reconstructed neutron count rate are driven by underlying changes in the hydrogen distribution is not a trivial task and the possibility remains that a more thorough examination of potential sites would yield a positive result. At the poles the change in neutron flux is dominated by changes in CO_2 frost cover. The reconstructions show that the seasonal cap forms away from the poles in both the north and south and contains significant structure. There is broad agreement between the CO_2 abundances implied by neutron data and the predictions of the GCMs, but at most latitudes less CO_2 is seen in the neutron data and the cap seems to persist further into the spring than expected given the results of the GCM.

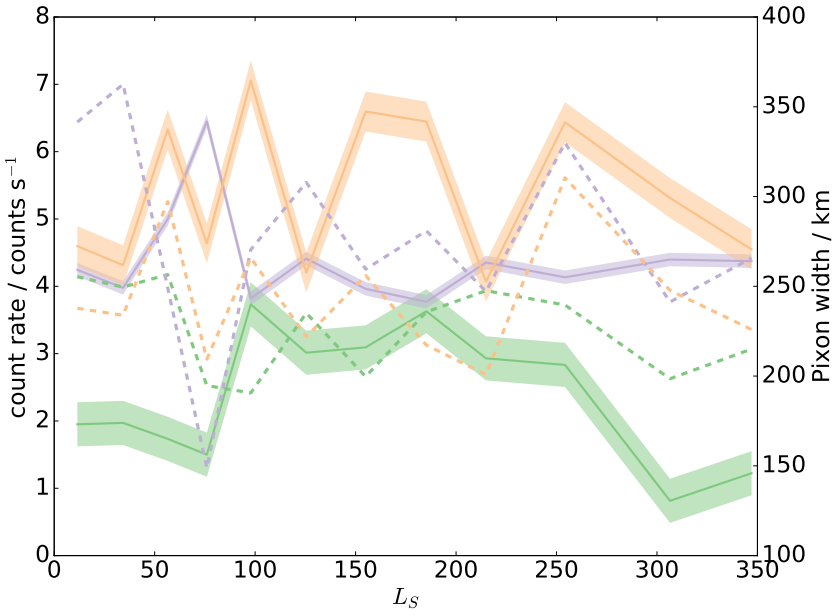

Figure 6.9: The count rate variation with Martian solar longitude for the three pixels at equatorial latitudes that have the largest fractional scatter in figure 6.8. The solid lines are for pixels at 1° S, 149° E (green), 32° S, 103° W (orange) and 12° N, 172° W (violet). The translucent regions indicate 1σ errors, determined empirically as the standard deviation between a series of reconstructions based on mock data (see section 4.3.2 for details). The dashed lines correspond to the pixon widths, in the same pixels, and are read from the right-hand y-axis.

CHAPTER 7

Conclusions

Bayesian image reconstruction is a powerful tool. It provides a method to remove blurring caused by the detector and suppress noise. To increase its utility in planetary science we have developed and implemented versions of the pixon image reconstruction technique (Pina & Puetter, 1993) on the sphere, enabling the reconstruction of globally-defined data sets. We have implemented, on the sphere, a form of the existing maximum entropy pixon method. Additionally, we developed and applied two new versions of the pixon method to address possible deficiencies in the maximum entropy method. Firstly, the hard pixon method, which explicitly uses discrete, hard-edged pixons to describe an image, thus enabling the creation of sharp boundaries in an image. Secondly, the locally adaptive pixon method that responds to localized misfit in the reconstruction, adding in structure only when it improves the fit locally, unlike the maximum entropy method wherein structure may be allowed in a region where it is not required. This new locally adaptive method is seen to closely approximate the maximum entropy method when there is constant noise throughout the image. The most productive approach to image reconstruction will depend on the data being reconstructed and hopefully each of these techniques will prove useful when applied to some global planetary data set.

In this thesis we have applied these image reconstruction techniques to each of a remotely sensed gamma ray and neutron data set. Examination of the Th-line LP-GRS data around the Compton-Belkovich Volcanic Complex yielded the highest resolution map of the lunar Th distribution to date. Combining this new data set with imaging revealed that pyroclastic eruption at the formation of the complex spread Th-rich material up to 40 km beyond the complex itself. This observation was shown to be consistent with the pyroclastic eruption of silicic material from the volcanic complex during its formation, but not with post deposition transport away from the complex due to impact processes.

Application of the locally adaptive pixon reconstruction technique to the entire global Mars Odyssey epithermal neutron data yielded a near-two-fold improvement in the linear spatial resolution of WEH maps, taking the resolution from 520 to 290 km. The suppression of noise due to image reconstruction allows the identification of features that would, in the raw data, not be discernible. These new maps revealed WEH-rich features close to the equator at the Medusae Fossae Formation and Tharsis Montes, allowing the likely existence of near-surface water ice rich material to be inferred. Conversely, no evidence for increased hydration is seen at the plate like features suggested to be a buried water ice sea (Murray et al., 2005) and the WEH content at the locations containing RSL is not significantly different, on scales larger than 120 km, from those without.

Splitting the MONS data by solar longitude we have also looked for seasonal changes in the Martian epithermal neutron flux and found that the deposition and sublimation of the Martian polar, seasonal CO_2 caps is a non-uniform process, with the caps offset from the poles. The approximate timing of seasonal cap deposition and sublimation seen in the MONS data is in agreement with Martian GCMs, but total amounts of deposited CO_2 differ by up to 50%. We see no evidence for significant variation in epithermal neutron flux away from the poles.

7.1 Ongoing and future work

The new image reconstruction techniques that have been developed in this thesis should provide ways to improve the utility of any data set defined on the sphere, allowing the valuable data returned by remote sensing missions to be explored fully. We have considered the Lunar Prospector Th-line gamma ray data around the CBVC and derived interesting geological results. As image reconstruction typical extends the dynamic range of a signal, performing a reconstruction of this data covering the entire lunar surface may reveal further compositionally interesting units that have not been previously identified. Additionally this data could be used to examine the variation of the lunar crustal composition with depth. The Moons crust is believed to be the result of the differential crystallization of a global magma ocean. One prediction of this hypothesis is that the composition of the pristine crust should vary with depth. The lower parts of the crust form later so

become progressively more mafic and enriched in incompatible elements. Measurement of this compositional variation, without direct access to samples is a difficult problem that has, so far, relied only on visible and infra red spectroscopic data because of the relatively poor spatial resolution of the neutron and gamma ray data. However, combining a pixon reconstruction of the neutron and gamma ray data with crustal thickness measurements, such as derived from GRAIL data, will enable the bulk composition of the crust to be determined, which, in turn, places constraints on the extent of differentiation and the ultimate source of the crustal material.

In this thesis we have considered the data to be either static in time, by stacking all observations where CO_2 frost is not present, or seasonally varying. However, as Odyssey has now been in orbit for over seven Mars years, it is possible use the neutron data to look for secular evolution of the Martian surface. Jakosky & Haberle (1990) suggested that the south pole residual cap may be unstable from year-to-year, with disappearance and reappearance controlled by small variations in climate. More recently, Colaprete et al. (2005) used a GCM to show that, even without changes in insolation, the SPRC can shift between regimes of losing and gaining mass due to albedo changes from increased dust loading or water ice cover. Additionally, the expansion of the 'swiss cheese' polar ice features reported by Malin et al. (2001) and Byrne & Ingersoll (2003) are consistent with the notion of a SPRC that is only quasi-stable and may, currently, be undergoing a reduction in size. With seven years of data we could look for evolution in the size of the epithermal neutron flux peak over the residual cap.

Another possible area of study is looking at the correlation of seasonal polar cap mass with global dust activity. Viking pressure measurements suggest that the CO_2 cycle is fairly consistent between years. However, airborne dust is known to influence the formation of the seasonal CO_2 cap (Bonev et al., 2008) and is variable year-on-year. Models have been made that attempt to reconcile these facts (Kahre & Haberle, 2010), but with seven years of neutron data we could look at precisely at how dust storms influence the timing of deposition and sublimation of the seasonal caps, by correlating these times as seen with the neutron data with known dust storm activity, along with the formation of any inhomogeneities within the caps, which will be made apparent using our new image reconstruction techniques.

The hard pixon method, developed in chapter 2, has not be used to reconstruct gamma ray or neutron data in this thesis due to its failure to adequately suppress noise in the test data. However, it was seen to provide the reconstruction that most closely matched the true test image. This suggests that the method has potential. Further development and optimisation of this algorithm, including incorporation of the spatial structure of the pixons into the prior, should yield a useful tool.

7.2 Final remarks

Throughout this thesis we have been concerned with determining the elemental composition of the Moon and Mars. We have derived new image reconstruction techniques that are defined on the sphere and used them to produce improved, higher-resolution maps of elemental abundance. These new techniques, and the maps that they afford, provide an upgraded resource for those studying planetary surface composition and the techniques can be used to improve any present, or future, data set defined on the sphere with a well understood PSF and noise distribution.

Bibliography

Albee, A. L., Arvidson, R. E., Palluconi, F., & Thorpe, T. 2001, Journal of Geophysical Research (Planets), 106, 23291

Anderson, H. R. 1965, Science, 149, 1226

Appéré, T., Schmitt, B., Langevin, Y., et al. 2011, Journal of Geophysical Research (Planets), 116, E05001

Arvidson, R., Drozd, R. J., Hohenberg, C. M., Morgan, C. J., & Poupeau, G. 1975, Moon, 13, 67

Baldwin, R. B. 1949, The face of the moon. (Univ. of Chicago Press)

Bandfield, J. L., Hamilton, V. E., & Christensen, P. R. 2000, Science, 287, 1626

Barghouty, A. F., Meyer, F. W., Harris, P. R., & Adams, J. H. 2011, Nuclear Instruments and Methods in Physics Research B, 269, 1310

Barnes, J. R., Pollack, J. B., Haberle, R. M., et al. 1993, Journal of Geophysical Research (Planets), 98, 3125

Berman, D. C., & Hartmann, W. K. 2002, Icarus, 159, 1

Bhandari, N. 2005, Journal of Earth System Science, 114, 701

Bhattacharya, S., Saran, S., Dagar, A., et al. 2013, Current Science, 105, 685

Bibring, J.-P., Langevin, Y., Mustard, J. F., et al. 2006, Science, 312, 400

Binder, A. B. 1998, Science, 281, 1475

Blewett, D. T., Lucey, P. G., Hawke, B. R., & Jolliff, B. L. 1997, Journal of Geophysical Research (Planets), 102, 16319

Bonev, B. P., Hansen, G. B., Glenar, D. A., James, P. B., & Bjorkman, J. E. 2008, Planetary and Space Science, 56, 181

Boynton, W., Feldman, W., Mitrofanov, I., et al. 2004, Space Science Reviews, 110, 37

Bracewell, R. N. 1958, Proceedings of the Institute of Radio Engineers, 46, 97

Brown, A. J., Calvin, W. M., McGuire, P. C., & Murchie, S. L. 2010, Journal of Geophysical Research (Planets), 115, E00D13

Brown, A. J., Calvin, W. M., & Murchie, S. L. 2012, Journal of Geophysical Research (Planets), 117, E00J20

Burnham, D. L. 1993, Spaceflight, 35, 278

Burr, D. M., Grier, J. A., McEwen, A. S., & Keszthelyi, L. P. 2002, Icarus, 159, 53

Byers, B. 1977, Destination Moon: A History of the Lunar Orbiter Program, NASA technical memorandum (National Aeronautics and Space Administration)

Byrne, S., & Ingersoll, A. P. 2003, Geophysical Research Letters, 30, 1696

Campbell, B. A., Putzig, N. E., Carter, L. M., et al. 2013, Journal of Geophysical Research (Planets), 118, 436

Carr, M. H. 1996, Water on Mars

Cassidy, T. A., & Johnson, R. E. 2005, Icarus, 176, 499

Cassini, G. D. 1666, Philosophical Transactions of the Royal Society of London Series I, 2, 615

Castro, J. M., Schipper, C. I., Mueller, S. P., et al. 2013, Bulletin of Volcanology, 75, 702

Chevrel, S. D., Pinet, P. C., & Head, J. W. 1999, Journal of Geophysical Research (Planets), 104, 16515

Chevrier, V. F., Hanley, J., & Altheide, T. S. 2009, Geophys. Res. Lett., 36, 10202

Christensen, P. R., Morris, R. V., Lane, M. D., Bandfield, J. L., & Malin, M. C. 2001, Journal of Geophysical Research (Planets), 106, 23873

Christensen, P. R., Jakosky, B. M., Kieffer, H. H., et al. 2004, Space Sci. Rev., 110, 85

Clancy, R. T., Sandor, B. J., Wolff, M. J., et al. 2012, Journal of Geophysical Research (Planets), 117, E00J10

Clifford, S. M., Lasue, J., Le Gall, A., & Heggy, E. 2011, in Lunar and Planetary Science Conference, Vol. 42, Lunar and Planetary Science Conference, 2142

Colaprete, A., Barnes, J. R., Haberle, R. M., et al. 2005, Nature, 435, 184

Daydou, Y. H., Pinet, P. C., Chevrel, S., & Le Mouélic, S. 2003, Planetary and Space Science, 51, 309

Driscoll, J. R., & Healy, D. M. 1994, Advances in Applied Mathematics, 15, 202

Dundas, C. M., Byrne, S., McEwen, A. S., et al. 2014, Journal of Geophysical Research (Planets), 119, 109

Dundas, C. M., & McEwen, A. S. 2015, Icarus, 254, 213

Edgett, K. S., Williams, R. M. E., Malin, M. C., Cantor, B. A., & Thomas, P. C. 2003, Geomorphology, 52, 289

Eke, V. 2001, MNRAS, 324, 108

Eke, V. R., Teodoro, L. F. A., & Elphic, R. C. 2009, Icarus, 200, 12

Elphic, R. C., Feldman, W. C., Prettyman, T. H., et al. 2005a, in Lunar and Planetary Science Conference, Vol. 36, 36th Annual Lunar and Planetary Science Conference, ed. S. Mackwell & E. Stansbery

Elphic, R. C., Lawrence, D. J., Feldman, W. C., et al. 1998, Science, 281, 1493

—. 2000, Journal of Geophysical Research (Planets), 105, 20333

Elphic, R. C., Lawrence, D. J., Feldman, W. C., et al. 2005b, in Lunar and Planetary Science Conference, Vol. 36, 36th Annual Lunar and Planetary Science Conference, ed. S. Mackwell & E. Stansbery

Elphic, R. C., Feldman, W. C., Prettyman, T. H., et al. 2004, in Lunar and Planetary Science Conference, Vol. 35, Lunar and Planetary Science Conference, ed. S. Mackwell & E. Stansbery

Fastook, J. L., Head, J. W., Marchant, D. R., & Forget, F. 2008, Icarus, 198, 305

Feldman, W. C., Barraclough, B. L., Fuller, K. R., et al. 1999, Nuclear Instruments and Methods in Physics Research A, 422, 562

Feldman, W. C., Barraclough, B. L., Maurice, S., et al. 1998, Science, 281, 1489

Feldman, W. C., & Drake, D. M. 1986, Nuclear Instruments and Methods in Physics Research A, 245, 182

Feldman, W. C., Gasnault, O., Maurice, S., et al. 2002, Journal of Geophysical Research (Planets), 107, 5016

Feldman, W. C., Lawrence, D. J., Elphic, R. C., et al. 2000, Journal of Geophysical Research (Planets), 105, 20347

Feldman, W. C., Pathare, A., Maurice, S., et al. 2011, Journal of Geophysical Research (Planets), 116, E11009

Feldman, W. C., Maurice, S., Lawrence, D. J., et al. 2001, Journal of Geophysical Research (Planets), 106, 23231

Feldman, W. C., Prettyman, T. H., Boynton, W. V., et al. 2003, Journal of Geophysical Research (Planets), 108, 7

Feldman, W. C., Prettyman, T. H., Maurice, S., et al. 2004, Journal of Geophysical Research (Planets), 109, 9006

Fichtel, C. E., & Trombka, J. I. 1997, Gamma-ray astrophysics. New insight into the universe. Second edition.

Fogel, R. A., & Rutherford, M. J. 1995, Geochimica et Cosmochimica Acta, 59, 201

Folch, A., & Felpeto, A. 2005, Journal of Volcanology and Geothermal Research, 145, 337

Forget, F., Haberle, R. M., Montmessin, F., Levrard, B., & Head, J. W. 2006, Science, 311, 368

Forget, F., Hourdin, F., Fournier, R., et al. 1999, Journal of Geophysical Research (Planets), 104, 24155

Fournier, G. 1914, L'Astronomie, 28, 17

Gaisser, T. K. 1990, Cosmic rays and particle physics

Galilei, G. 1610, Sidereus nuncius magna, longeque admirabilia spectacula pandens lunae facie, fixis innumeris, lacteo circulo, stellis nebulosis, ... Galileo Galileo : nuper a se reperti beneficio sunt observata in apprime vero in quatuor planetis circa Iovis stellam disparibus intervallis, atque periodis, celeritate mirabili circumvolutis ... atque Medicea sidera nuncupandos decrevit

Garrett, A. J. M. 2001, in American Institute of Physics Conference Series, Vol. 568, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. A. Mohammad-Djafari, 3–22

Gasnault, O., Feldman, W. C., Maurice, S., et al. 2001, Geophys. Res. Lett., 28, 3797

Gault, D. E., Hoerz, F., Brownlee, D. E., & Hartung, J. B. 1974, in Lunar and Planetary Science Conference Proceedings, Vol. 5, Lunar and Planetary Science Conference Proceedings, 2365–2386

Gendrin, A., Mangold, N., Bibring, J.-P., et al. 2005, Science, 307, 1587

Gilbert, G. K. 1893, The moon's face. A study of the origin of its features.

Gilbert, W. 1651, De Mundo nostro sublunari philosophia nova

Gillis, J. J., Jolliff, B. L., Lawrence, D. J., Lawson, S. L., & Prettyman, T. H. 2002, in Lunar and Planetary Institute Science Conference Abstracts, Vol. 33, Lunar and Planetary Institute Science Conference Abstracts, 1967

Glavin, D. P., Freissinet, C., Miller, K. E., et al. 2013, Journal of Geophysical Research (Planets), 118, 1955

Glotch, T. D., Lucey, P. G., Bandfield, J. L., et al. 2010, Science, 329, 1510

Goldstein, D. B., Nerem, R. S., Barker, E. S., et al. 1999, Geophys. Res. Lett., 26, 1653

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

Greeley, R. 1986, in LPI Contributions, Vol. 599, Mars: Evolution of its Climate and Atmosphere, 29

Greeley, R., & Guest, J. E. 1987, Geologic map of the eastern equatorial region of Mars, Tech. rep.

Grimm, R. E., Harrison, K. P., & Stillman, D. E. 2014, Icarus, 233, 316

Gwinner, K., Jaumann, R., Hauber, E., et al. 2016, Planetary and Space Science, 126, 93

Haberle, R. M., Mattingly, B., & Titus, T. N. 2004, Geophys. Res. Lett., 31, L05702

Haberle, R. M., Pollack, J. B., Barnes, J. R., et al. 1993, Journal of Geophysical Research (Planets), 98, 3093

Hagerty, J. J., Lawrence, D. J., Hawke, B. R., & Gaddis, L. R. 2009, Journal of Geophysical Research (Planets), 114, 4002

Hagerty, J. J., Lawrence, D. J., Hawke, B. R., et al. 2006, Journal of Geophysical Research (Planets), 111, 6002

Harri, A.-M., Genzer, M., Kemppinen, O., et al. 2014, Journal of Geophysical Research: Planets, 119, 82

Harrison, S., Balme, M., Hagermann, A., Murray, J., & Muller, J.-P. 2010, Icarus, 209, 405

Hawke, B. R., Lawrence, D. J., Blewett, D. T., et al. 2003, Journal of Geophysical Research (Planets), 108, 5069

Hayward, R., Fenton, L., Tanaka, K., et al. 2010, Mars Global Digital Dune Database, Tech. rep.

Head, J. W. 1976, Reviews of Geophysics and Space Physics, 14, 265

Head, J. W., & Kreslavsky, M. 2004, in Lunar and Planetary Science Conference, Vol. 35, Lunar and Planetary Science Conference, ed. S. Mackwell & E. Stansbery

Head, J. W., & Marchant, D. R. 2003, Geology, 31, 641

Head, J. W., & Weiss, D. K. 2014, Planetary and Space Science, 103, 331

Head, J. W., Wilson, L., Robinson, M., et al. 2000, in Environmental Effects on Volcanic Eruptions: From Deep Oceans to Deep Space, ed. J. R. Zimbelman & T. K. P. Gregg, 143

Head, J. W., Wilson, L., & Weitz, C. M. 2002, Journal of Geophysical Research (Planets), 107, 5001

Hecht, M. H., Kounaves, S. P., Quinn, R. C., et al. 2009, Science, 325, 64

Heiken, G., Vaniman, D., & French, B. 1991, Lunar Sourcebook: A User's Guide to the Moon (Cambridge University Press), 357–374

Helstrom, C. W. 1967, Journal of the Optical Society of America (1917-1983), 57, 297

Herschel, W. 1784, Philosophical Transactions of the Royal Society of London Series I, 74, 233

Hess, S. L., Henry, R. M., & Tillman, J. E. 1979, Journal of Geophysical Research (Planets), 84, 2923

Hess, S. L., Ryan, J. A., Tillman, J. E., Henry, R. M., & Leovy, C. B. 1980, Geophys. Res. Lett., 7, 197

Hörz, F. 1977, Physics and Chemistry of Earth, 10, 3

Housen, K. R., Schmidt, R. M., & Holsapple, K. A. 1983, Journal of Geophysical Research (Planets), 88, 2485

Houston, W. N., Moriwaki, Y., & Chang, C.-S. 1973, in Lunar and Planetary Science Conference Proceedings, Vol. 4, Lunar and Planetary Science Conference Proceedings, 2425

Huygens, C. 1698, [Kosmotheoros]; sive, de terris coelestibus earumque ornatu conjecturae.

Jakosky, B. M., & Haberle, R. M. 1990, Journal of Geophysical Research (Planets), 95, 1359

James, P. B., Kieffer, H. H., & Paige, D. A. 1992, The seasonal cycle of carbon dioxide on Mars, ed. M. George, 934–968

Jansson, P. A. 1984, Deconvolution. With applications in spectroscopy (Academic Press, Inc.)

Jolliff, B. L. 1998, International Geology Review, 40, 916

Jolliff, B. L., Gillis, J. J., Haskin, L. A., Korotev, R. L., & Wieczorek, M. A. 2000, Journal of Geophysical Research (Planets), 105, 4197

Jolliff, B. L., Lawrence, S. J., Robinson, M. S., et al. 2011a, LPI Contributions, 1646, 32

Jolliff, B. L., Wiseman, S. A., Lawrence, S. J., et al. 2011b, Nature Geoscience, 4, 566

Kahre, M. A., & Haberle, R. M. 2010, Icarus, 207, 648

Kato, M., Sasaki, S., Tanaka, K., Iijima, Y., & Takizawa, Y. 2008, Advances in Space Research, 42, 294

Kelly, N. J., Boynton, W. V., Kerry, K., et al. 2006, Journal of Geophysical Research (Planets), 111, E03S07

Kieffer, H. H. 1979, Journal of Geophysical Research (Planets), 84, 8263

Kieffer, H. H., Jakosky, B. M., Snyder, C. W., & Matthews, M. S. 1992, Mars

Kieffer, H. H., & Titus, T. N. 2001, Icarus, 154, 162

Kieffer, S. W. 1981, in The 1980 eruptions of Mount St. Helens, ed. P. Lipman & D. Mullineaux (US Geol. Surv. Prof. Pap), 379–400

Kieffer, S. W., & Sturtevant, B. 1984, Journal of Geophysical Research: Solid Earth (1978–2012), 89, 8253

Kuiper, G. P. 1955, Publications of the Astronomical Society of the Pacific, 67, 271

—. 1957, ApJ, 125, 307

Lancaster, N., & Greeley, R. 1990, Journal of Geophysical Research (Planets), 95, 10921

Langner, U. W., Potgieter, M. S., & Webber, W. R. 2003, Journal of Geophysical Research: Space Physics, 108, 8039

Lawrence, D. J., Elphic, R. C., Feldman, W. C., Funsten, H. O., & Prettyman, T. H. 2010, Astrobiology, 10, 183

Lawrence, D. J., Elphic, R. C., Feldman, W. C., et al. 2003, Journal of Geophysical Research (Planets), 108, 5102

Lawrence, D. J., Feldman, W. C., Barraclough, B. L., et al. 1999, Geophys. Res. Lett., 26, 2681

—. 2000, Journal of Geophysical Research (Planets), 105, 20307

—. 1998a, Science, 281, 1484

—. 1998b, Science, 281, 1484

Lawrence, D. J., Feldman, W. C., Elphic, R. C., et al. 2002, Journal of Geophysical Research (Planets), 107, 5130

Lawrence, D. J., Maurice, S., & Feldman, W. C. 2004, Journal of Geophysical Research (Planets), 109, 7

Lawrence, D. J., Puetter, R. C., Elphic, R. C., et al. 2007, Geophys. Res. Lett., 34, 3201

Le Mouélic, S., Langevin, Y., & Erard, S. 1999, Journal of Geophysical Research (Planets), 104, 3833

Leo, W. R. 1987, Techniques for Nuclear and Particle Physics Experiments: A How to Approach (Springer-Verlag, Berlin, Heidelberg)

Leovy, C., & Mintz, Y. 1969, Journal of Atmospheric Sciences, 26, 1167

Li, L., & Mustard, J. F. 2005, Journal of Geophysical Research (Planets), 110, 11002

Lingenfelter, R. E., Canfield, E. H., & Hess, W. N. 1961, Journal of Geophysical Research (Planets), 66, 2665

Litvak, M. L., Mitrofanov, I. G., Kozyrev, A. S., et al. 2006, Icarus, 180, 23

—. 2007, Journal of Geophysical Research (Planets), 112, E03S13

Lockwood, J. A. 1973, Space Sci. Rev., 14, 663

Lucey, P., Korotev, R. L., Gillis, J. J., et al. 2006, Reviews in Mineralogy and Geochemistry, 60, 83

Madeleine, J.-B., Forget, F., Head, J. W., et al. 2009, Icarus, 203, 390

Malin, M. C., Caplinger, M. A., & Davis, S. D. 2001, Science, 294, 2146

Malin, M. C., Edgett, K. S., Cantor, B. A., et al. 2010, International Journal of Mars Science and Exploration, 5, 1

Marcus, A. H. 1970, Moon, 1, 297

Maurice, S., Feldman, W., Diez, B., et al. 2011, Journal of Geophysical Research (Planets), 116, 11008

Maurice, S., Feldman, W. C., Lawrence, D. J., et al. 2000, Journal of Geophysical Research (Planets), 105, 20365

Maurice, S., Lawrence, D. J., Feldman, W. C., Elphic, R. C., & Gasnault, O. 2004, Journal of Geophysical Research (Planets), 109, E07S04

McEwen, A. S., Robinson, M. S., Eliason, E. M., et al. 1994, Science, 266, 1858

McEwen, A. S., Ojha, L., Dundas, C. M., et al. 2011, Science, 333, 740

McEwen, A. S., Dundas, C. M., Mattson, S. S., et al. 2014, Nature Geosci, 7, 53

Mellon, M. T., Jakosky, B. M., & Postawko, S. E. 1997, Journal of Geophysical Research (Planets), 102, 19357

Metzger, A. E., Haines, E. L., Parker, R. E., & Radocinski, R. G. 1977, in Lunar and Planetary Science Conference Proceedings, Vol. 8, Lunar and Planetary Science Conference Proceedings, ed. R. B. Merril, 949–999

Mitrofanov, I., Anfimov, D., Kozyrev, A., et al. 2002, Science, 297, 78

Mitrofanov, I. G., Litvak, M. L., Varenikov, A. B., et al. 2012, Space Sci. Rev., 170, 559

Mitrofanov, I. G., Litvak, M. L., Sanin, A. B., et al. 2014, Journal of Geophysical Research (Planets), 119, 1579

Mouginot, J., Pommerol, A., Beck, P., Kofman, W., & Clifford, S. M. 2012, Geophys. Res. Lett., 39, L02202

Mount, C. P., & Titus, T. N. 2015, Journal of Geophysical Research (Planets), 120, 1252

Murray, J. B., Muller, J.-P., Neukum, G., et al. 2005, Nature, 434, 352

Navarro, T., Madeleine, J.-B., Forget, F., et al. 2014, Journal of Geophysical Research (Planets), 119, 1479

Nielsen, E. 2004, Space Science Reviews, 111, 245

Nolan, J. P. 1999, in Practical Guide to Heavey Tails: Statistical Techniques and Applications, ed. R. J. Alder, R. E. Feldman, & M. S. Taqqu, 527–533

Nozette, S., Rustan, P., Pleasance, L. P., et al. 1994, Science, 266, 1835

Ojha, L., Wilhelm, M. B., Murchie, S. L., et al. 2015, Nature Geosci, 8, 829

Ouyang, Z., Jiang, J., Li, C., et al. 2008, Chin. J. Space Sci, 28, 361

Paige, D. A., Bachman, J. E., & Keegan, K. D. 1994, Journal of Geophysical Research (Planets), 99, 25959

Parker, E. N. 1965, Planetary and Space Science, 13, 9

Parsons, R. L., & Head, J. W. 2005, in Lunar and Planetary Science Conference, Vol. 36, 36th Annual Lunar and Planetary Science Conference, ed. S. Mackwell & E. Stansbery

Pelowitz, D. B. 2005, MCNPX Users Manual Version 2.5.0 (Los Alamos Natl. Lab., Los Alamos, N. M.: Los Alamos Natl. Lab.)

Petro, N. E., Isaacson, P. J., Pieters, C. M., et al. 2013, in Lunar and Planetary Inst. Technical Report, Vol. 44, Lunar and Planetary Science Conference, 2688

Phillips, R. J., Zuber, M. T., Smrekar, S. E., et al. 2008, Science, 320, 1182

Picardi, G., Biccari, D., Seu, R., et al. 2004, Planetary and Space Science, 52, 149

Picardi, G., Plaut, J. J., Biccari, D., et al. 2005, Science, 310, 1925

Pieters, C. M., & Englert, P. A. J. 1993, Remote Geochemical Analysis, Elemental and Mineralogical Composition

Pina, R. K., & Puetter, R. C. 1992, Publications of the Astronomical Society of the Pacific, 104, 1096

—. 1993, Publications of the Astronomical Society of the Pacific, 105, 630

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in C. The art of scientific computing

Prettyman, T. H. 2007, Remote Chemical Sensing Using Nuclear Spetroscopy, ed. L.-A. A. McFadden, P. R. Weissman, & T. V. Johnson, 765–786

Prettyman, T. H., Feldman, W. C., & Titus, T. N. 2009, Journal of Geophysical Research (Planets), 114, 8005

Prettyman, T. H., Hagerty, J. J., Elphic, R. C., et al. 2006, Journal of Geophysical Research (Planets), 111, 12007

Prettyman, T. H., Lawrence, D. J., Vaniman, D. T., Elphic, R. C., & Feldman, W. C. 2002a, in The Moon Beyond 2002: Next Steps in Lunar Science and Exploration, 49

Prettyman, T. H., Feldman, W. C., Lawrence, D. J., et al. 2002b, in Lunar and Planetary Institute Science Conference Abstracts, Vol. 33, Lunar and Planetary Institute Science Conference Abstracts, 2012

Prettyman, T. H., Feldman, W. C., Mellon, M. T., et al. 2004, Journal of Geophysical Research (Planets), 109, E05001

Prettyman, T. H., Feldman, W. C., McSween, H. Y., et al. 2011, Space Sci. Rev., 163, 371

Puetter, R. C. 1996, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2827, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. P. S. Idell & T. J. Schulz, 12–31

Reedy, R. C. 1978, in Lunar and Planetary Science Conference Proceedings, Vol. 9, Lunar and Planetary Science Conference Proceedings, 2961–2984

Reedy, R. C., Arnold, J. R., & Trombka, J. I. 1973, Journal of Geophysical Research (Planets), 78, 5847

Reinecke, M., & Seljebotn, D. S. 2013, Astronomy and Astrophysics, 554, A112

Righter, K., & O'Brien, D. P. 2011, Proceedings of the National Academy of Sciences, 108, 19165

Riley, K. F., Hobson, M. P., & Bence, S. J. 2006, Mathematical Methods for Physics and Engineering Third Edition Paperback Set, 1362

Risbo, T. 1996, Journal of Geodesy, 70, 383

Robinson, M., & Riner, M. 2005, Journal of Earth System Science, 114, 669

Robinson, M. S., Malaret, E., & White, T. 2003, Journal of Geophysical Research (Planets), 108, 9

Sagan, C., & Pollack, J. B. 1969, Nature, 223, 791

Salisbury, J. W., Vincent, R. K., Logan, L. M., & Hunt, G. R. 1970, Journal of Geophysical Research (Planets), 75, 2671

Saunders, R. S., & Blewett, D. T. 1987, Astronomicheskii Vestnik, 21, 181

Schröter, J. H. 1791, Selenotopographische Fragmente zur genauern Kenntniss der Atmosphäre, sammt den dazu gehörigen Specialcharten und Mondfläche, ihrer erlittenen Veränderungen und Zeichnungen

Schultz, P., & Lutz, A. B. 1988, Icarus, 73, 91

Scott, D. H., & Tanaka, K. L. 1982, Journal of Geophysical Research (Planets), 87, 1179

Scott, D. H., & Tanaka, K. L. 1986, Geologic map of the western equatorial region of Mars, Tech. rep.

Seddio, S. M., Jolliff, B. L., Korotev, R. L., & Zeigler, R. A. 2013, American Mineralogist, 98, 1697(17)

Seu, R., Phillips, R. J., Biccari, D., et al. 2007, Journal of Geophysical Research: Planets, 112, n/a, e05S05

Shean, D. E., Head, J. W., Fastook, J. L., & Marchant, D. R. 2007, Journal of Geophysical Research (Planets), 112, E03004

Shean, D. E., Head, J. W., & Marchant, D. R. 2005, Journal of Geophysical Research (Planets), 110, E05001

Shearer, C. K., Hess, P. C., Wieczorek, M. A., et al. 2006, Reviews in Mineralogy and Geochemistry, 60, 365

Shelton, W. 1968, Soviet space exploration: the first decade (Washington Square Press)

Shirley, K. A., Zanetti, M., Jolliff, B., van der Bogert, C. H., & Hiesinger, H. 2013, in Lunar and Planetary Institute Science Conference Abstracts, Vol. 44, Lunar and Planetary Institute Science Conference Abstracts, 2469

Shkuratov, Y. G., Kaidash, V. G., Kreslavsky, M. A., & Opanasenko, N. V. 2001, Solar System Research, 35, 29

Shoemaker, E. M. 1963, Impact Mechanics at Meteor Crater, Arizona, ed. G. P. Kuiper & B. M. Middlehurst, 301

Sides, I., Edmonds, M., Maclennan, J., et al. 2014, Earth and Planetary Science Letters, 400, 102

Smith, D. E., Zuber, M. T., & Neumann, G. A. 2001, Science, 294, 2141

Sparks, R. S. J., & Wilson, L. 1982, Geophysical Journal International, 69, 551

Spudis, P. D. 1999, The Moon, ed. J. K. Beatty, C. Collins Petersen, & A. Chaikin, 125

Spurr, J. E. 1944, Geology applied to selenology; the Imbrian plain region of the moon

Squyres, S. 2011, Vision and Voyagers for Planetary Science in the Decade 2013-2022 (The National Academies press, Washington DC)

Steele, L. J., Lewis, S. R., Patel, M. R., et al. 2014, Icarus, 237, 97

Stillman, D. E., Michaels, T. I., Grimm, R. E., & Hanley, J. 2016, Icarus, 265, 125

Stillman, D. E., Michaels, T. I., Grimm, R. E., & Harrison, K. P. 2014, Icarus, 233, 328

Tanaka, K., Skinner, J., & Hare, T. 2005, Geologic Map of the Northern Plains of Mars., Tech. rep.

Tanaka, K., Skinner, J., Dohm, J., et al. 2014, Geologic map of Mars: U.S. Geological Survey Scientific Investigations Map 3292, Tech. rep.

Taylor, F. W. 2010, The Scientific Exploration of Mars

Titus, T. N., Kieffer, H. H., & Christensen, P. R. 2003, Science, 299, 1048

Travis, B. J., Feldman, W. C., & Maurice, S. 2013, Journal of Geophysical Research (Planets), 118, 877

Uesugi, K. 1996, Advances in Space Research, 18, 69

van Cittert, P. H. 1931, Zeitschrift fur Physik, 69, 298

Vinogradov, A. P., Surkov, Y. A., Chernov, G. M., Kirnozov, F. F., & Nazarkina, G. B. 1966, Cosmic Research, 4, 751

Vondrak, R., Keller, J., Chin, G., & Garvin, J. 2010, Space Science Reviews, 150, 7

Wandelt, B. D., & Górski, K. M. 2001, Physical Review D, 63, 123002

Wang, A., & Ling, Z. C. 2011, Journal of Geophysical Research (Planets), 116, E00F17

Wapstra, A. H., & Audi, G. 1985, Nuclear Physics A, 432, 1

Warren, P. H., & Wasson, J. T. 1979, Reviews of Geophysics and Space Physics, 17, 73

Watters, T. R., Campbell, B., Carter, L., et al. 2007, Science, 318, 1125

Whitaker, E. 1999, Mapping and Naming the Moon

Wieczorek, M. A., Jolliff, B. L., Khan, A., et al. 2006, Reviews in Mineralogy and Geochemistry, 60, 221

Wilhelms, D. E., McCauley, J. F., & Trask, N. J. 1987, The geologic history of the moon

Wilson, J. T., Eke, V. R., Massey, R. J., et al. 2015, Journal of Geophysical Research (Planets), 120, 92

Wilson, L., & Head, J. W. 2003, Geophys. Res. Lett., 30, 1605

Wood, C. A. 1980, Journal of Volcanology and Geothermal Research, 7, 387

Wurz, P., Rohner, U., Whitby, J. A., et al. 2007, Icarus, 191, 486

Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, ApJ, 449, 446